

Gencore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 31, 2002, 14:53:30 ; Search time 12.56 Seconds

Perfect score: 1 MSVGFFIGAQQLAYRFTANGI.....TPSSPGKULTRSLAIGKKD 314

Scoring table: BLOSUM62 Gapop 10.0 , Gapext 0.5

Title: US-09-912-717-1
1567 1 MSVGFFIGAQQLAYRFTANGI.....TPSSPGKULTRSLAIGKKD 314
562.582 Million cell updates/sec

Searched: 212252 seqs, 22503292 residues

Total number of hits satisfying chosen parameters: 212252

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : issued_Patents_AA:*

1: /cgn2_6/ptodata/2/1aa/5A_COMB.pep: *
2: /cgn2_6/ptodata/2/1aa/5B_COMB.pep: *
3: /cgn2_6/ptodata/2/1aa/6A_COMB.pep: *
4: /cgn2_6/ptodata/2/1aa/6B_COMB.pep: *
5: /cgn2_6/ptodata/2/1aa/PCTRUS_COMB.pep: *
6: /cgn2_6/ptodata/2/1aa/backfiles1.pep: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query	Match	Length	DB	ID	Description
1	1567	100.0	314	3	US-09-099-676-1		Sequence 1, Appli
2	1567	100.0	314	4	US-09-099-676-1		Sequence 1, Appli
3	1318	84.1	315	3	US-09-099-676-3		Sequence 1, Appli
4	1318	84.1	315	4	US-09-099-676-3		Sequence 1, Appli
5	294	18.8	304	1	US-08-665-910-3		Sequence 1, Appli
6	102	6.5	349	3	US-08-665-910-3		Sequence 1, Appli
7	98	6.3	367	3	US-08-605-150A-6		Sequence 1, Appli
8	98	6.3	381	3	US-08-605-150A-4		Sequence 1, Appli
9	97.5	6.2	543	1	US-08-605-150A-4		Sequence 1, Appli
10	97.5	6.2	543	1	US-08-752-929-17		Sequence 1, Appli
11	97.5	6.2	543	4	US-09-090-793-10		Sequence 1, Appli
12	95.5	6.1	1455	2	US-08-726-012B-2		Sequence 1, Appli
13	94.5	6.0	372	3	US-08-605-150A-8		Sequence 1, Appli
14	94.5	6.0	372	3	US-08-605-150A-10		Sequence 1, Appli
15	91.5	5.8	1026	2	US-08-614-977A-7		Sequence 1, Appli
16	91.5	5.8	1026	4	US-09-142-648B-7		Sequence 1, Appli
17	88.5	5.6	722	4	US-08-961-083-84		Sequence 1, Appli
18	88	5.6	372	3	US-08-605-150A-2		Sequence 1, Appli
19	85	5.4	398	1	US-08-176-413-5		Sequence 1, Appli
20	85	5.4	398	5	PCT-US94-14919-5		Sequence 1, Appli
21	84	5.4	4545	2	US-08-804-227C-14		Sequence 1, Appli
22	84	5.4	4550	2	US-08-804-227C-8		Sequence 1, Appli
23	84	5.4	4550	2	US-08-804-198-2		Sequence 1, Appli
24	83.5	5.3	801	1	US-08-725-012-2		Sequence 1, Appli
25	83	5.3	657	1	US-08-264-534-34		Sequence 1, Appli
26	83	5.3	1	US-08-083-590A-13		Sequence 1, Appli	
27	83	5.3	657	1	US-08-465-500-34		Sequence 1, Appli

ALIGNMENTS

RESULT 1
US-09-912-717-1
Sequence 1, Application US-09099676
Patent No. 610075
GENERAL INFORMATION:
APPLICANT: Hillman, Jennifer L.
APPLICANT: Corley, Neil C.
APPLICANT: Baughn, Mariah R.
TITLE OF INVENTION: DELTA 1-PYRROLINE-5-CARBOXYLATE REDUCTASE
TITLE OF INVENTION: HOMOLOG
NUMBER OF SEQUENCES: 3
CORRESPONDENCE ADDRESS:
ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: Windows
SOFTWARE: FastSQL for Windows Version 2.0b
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/099,676
FILING DATE: HEREWITHE
CLASSIFICATION:
PRIORITY APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Cerone, Michael C.
REGISTRATION NUMBER: 39,132
REFERENCE/DOCKET NUMBER: PF-0532 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650-855-0555
TELEFAX: 650-855-0572
TELEX:
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 314 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: PROSHION01
CLONE: 2278458
US-09-099-676-1

INFORMATION FOR SEQ ID NO: 3:

SEQUENCE CHARACTERISTICS:

LENGTH: 315 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

IMMEDIATE SOURCE:

LIBRARY: GENBANK

CLONE: 189498

US-09-099-676-3

Query Match 84.1%; Score 1318; DB 3; Length 315;
Best Local Similarity 83.8%; Pred. No. 4.7e-136; Mismatches 22; Indels 2; Gaps 1;
Matches 263; Conservative 27; MisMatches 22; Indels 2; Gaps 1;

Qy 1 MSVFIGAGOLAFAGKFTAGVLAHKIMASSPDMOLATVALSRKMGVNLTNSKETVKH 58
Db 1 MSVFIGAGOLAFAGKFTAGVLAHKIMASSPDMOLATVALSRKMGVNLTNSKETVKH 58
Db 59 SDVLFIAVKPHIIPFIDEAGDVAQHIVVSCAGVIVSVEKKLMAFOPAKVIRCM 118
Db 61 SDVLFIAVKPHIIPFIDEAGDIDEHRHIVVSCAGVIVSIEKKLSAFRPAPRVRICM 120

Qy 119 NTPVWVQEGATVATGTHALVEDQOLLEQMSVSGCTEVEDDIDAVGLSGSPAYAF 178
Db 121 NTPVWVQEGATVATGTHAQNEDGRIMEQOLLSVTGCTEVEDDIDAVGLSGSPAYAF 180

Qy 179 MALDADGGVKGMLPPLRATOLGAQALLGAAKMLDSEOPCQLKDNVCSGGATIHALHF 238

Db 181 TALDADGGVKGMLPPLRATOLGAQALLGAAKMLDSEOPCQLKDNVCSGGATIHALHF 240

Qy 239 LESGGFRSLINAVAEASCIRTEQRLQSMADQEQVSPAAIKTILDKVKLDPAGTALSPSG 298

Db 241 LESGGFRSLINAVAEASCIRTEQRLQSMADQEQVSPAAIKTILDKVKLDPAGTALSPSG 300

Qy 299 PGKLITRSLAGK 312

Db 301 HTKLKPLRSLAPKG 314

RESULT 4

Sequence 3, Application US/09565910

Patent No. 6268192

GENERAL INFORMATION:

APPLICANT: Hillman, Jennifer L.

APPLICANT: Corley, Neil C.

APPLICANT: Baughn, Mariah R.

TITLE OF INVENTION: DELTA 1-PYRROLINE-5-CARBOXYLATE REDUCTASE

TITLE OF INVENTION: HOMOLOG

NUMBER OF SEQUENCES: 3

CORRESPONDENCE ADDRESS:

ADDRESSEE: Incyte Pharmaceuticals, Inc.

STREET: 3174 Porter Drive

CITY: Palo Alto

STATE: CA

COUNTRY: USA

ZIP: 94304

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible

OPERATING SYSTEM: Windows

SOFTWARE: FASTEQ for Windows Version 2.0b

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/565,910

FILING DATE: 09/09/1999

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 09/099, 676

FILING DATE: 23-JUN-1995

ATTORNEY/AGENT INFORMATION:

NAME: Cerrone, Michael C

REGISTRATION NUMBER: 39,132

REFERENCE/DOCKET NUMBER: PF-0532 US

TELECOMMUNICATION INFORMATION:

TELEPHONE: 650-855-0555

TELEFAX: 650-855-0572

TELEX:

INFORMATION FOR SEQ ID NO: 3:

SEQUENCE CHARACTERISTICS:

LENGTH: 315 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

IMMEDIATE SOURCE:

LIBRARY: GENBANK

CLONE: 189498

US-09-565-910-3

Query Match 84.1%; Score 1318; DB 4; Length 315;
Best Local Similarity 83.8%; Pred. No. 4.7e-136; Mismatches 22; Indels 2; Gaps 1;
Matches 263; Conservative 27; MisMatches 22; Indels 2; Gaps 1;

Qy 1 MSVFIGAGOLAFAGKFTAGVLAHKIMASSPDMOLATVALSRKMGVNLTNSKETVKH 58
Db 1 MSVFIGAGOLAFAGKFTAGVLAHKIMASSPDMOLATVALSRKMGVNLTNSKETVKH 58
Db 59 SDVLFIAVKPHIIPFIDEAGDVAQHIVVSCAGVIVSVEKKLMAFOPAKVIRCM 118
Db 61 SDVLFIAVKPHIIPFIDEAGDIDEHRHIVVSCAGVIVSIEKKLSAFRPAPRVRICM 120

Qy 119 NTPVWVQEGATVATGTHALVEDQOLLEQMSVSGCTEVEDDIDAVGLSGSPAYAF 178
Db 121 NTPVWVQEGATVATGTHAQNEDGRIMEQOLLSVTGCTEVEDDIDAVGLSGSPAYAF 180

Qy 179 MALDADGGVKGMLPPLRATOLGAQALLGAAKMLDSEOPCQLKDNVCSGGATIHALHF 238

Db 181 TALDADGGVKGMLPPLRATOLGAQALLGAAKMLDSEOPCQLKDNVCSGGATIHALHF 240

Qy 239 LESGGFRSLINAVAEASCIRTEQRLQSMADQEQVSPAAIKTILDKVKLDPAGTALSPSG 298

Db 241 LESGGFRSLINAVAEASCIRTEQRLQSMADQEQVSPAAIKTILDKVKLDPAGTALSPSG 300

Qy 299 PGKLITRSLAGK 312

Db 301 HTKLKPLRSLAPKG 314

RESULT 5

Sequence 2, Application US/08665716

Patent No. 579222

GENERAL INFORMATION:

APPLICANT: KELLY, ROSEMARIE

APPLICANT: REGISTER, ELIZABETH A

APPLICANT: MASUREKAR, PRAKASH S

APPLICANT: ARBORICOLA

TITLE OF INVENTION: P5C REDUCTASE GENE FROM ZALERION

TITLE OF INVENTION: ARBORICOLA

NUMBER OF SEQUENCES: 2

CORRESPONDENCE ADDRESS:

ADDRESSEE: MERRCK & CO., INC.

STREET: 126 E. LINCOLN AVENUE

CITY: Rahway

STATE: NEW JERSEY

COUNTRY: US

ZIP: 07065-0900

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/665,716

FILING DATE: 23-JUN-1995

CLASSIFICATION:

ATTORNEY/AGENT INFORMATION:

NAME: KORSEN, ELLIOTT

REGISTRATION NUMBER: 32,705

REFERENCE/DOCKET NUMBER: 19453PV

TELECOMMUNICATION INFORMATION:

TELEPHONE: 908-594-5493

TELEFAX: 908-594-4720

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 304 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-665-716-2

Query Match 18.8%; Score 294; DB 1; Length 304;

Best Local Similarity 29.9%; Pred. No. 6.2e-24; Indels 38; Gaps 9; Matches 88; Conservative 49; Mismatches 119.

OY 1 MSVGFITGASOLARYTAAAGLISAKHITIASSPENN-----LPT--VSALR----- 42
Db 11 LTAVIGCGTMGIALISGIGAISIDEHAPNSOSSETDEPSKALPTKFACVRSRPGAEKI 70
OY 43 -----KMGVNLRNSRNKETV-KISDVILTLAVKHHIPFLDIEG-ADVOARHIVVSCAA 93
Db 71 KKALSPYKTPVKVITQSDNTVIACTREADAVLIGLCKPQYMAEGILGREGMVDAKGKLLISTIA 130
OY 94 GVTISVEKKLMAQOP-----KVRGTMTPVWVQEGATVYATCOTHAL-VEDGQL 145
Db 131 GVPABCGIYGYMGKTPVNPKEGLCQVVRAMPNTASGTRESMTVATSSPPLATSSLI 190
OY 146 EQIMSSVGFCTEVEEDLIDAVTGLSGSGPQAYAFAHMLDA-DGSKVKGMLPRLAQLGQA 203
Db 191 TWIFRIGDQVLPATMADSTALGGSGPAAEALTEADGAATAGLIPRAEORMAQAT 250
Db 251 MKGAGAGIVLVSCE-HFALLKDVKVTPGCCITGGLAVLLEGQVRSIVARAVREATV 303

Query Match 6.5%; Score 102; DB 3; Length 349;

Best Local Similarity 23.3%; Pred. No. 0.0083; Indels 60; Gaps 11; Matches 59; Conservative 39; Mismatches 95.

OY 59 SIVLFLAVKPHIIPFLDEIGADVOARHIVVSCAGYTISVSEKKLMAQOPPKVIRCM 118
Db 86 ADILWVYVPHQICDQKQDGLKLNKNTGSLINGVDEEFLNLIS-----EVIGERL 140
OY 119 NTPVVVVOEGATVYATGTHALVED-----GQLEQIMSSVGF-CTEVEE--- 160
Db 141 GIPMSVLMGANI---ASEVAEKFCTTIGCKDPAQGQQLKDQMOTPFRITVVOEVDT 196
OY 161 -DLIDAVTGLSGSGPQAYAFAHMLADAGGVKUMLPRLA-TDQGAQALLGAAKMLDSEQH 218
Db 197 VEIGGALKNNIVAVGAGFC-----DCIGFSNTKAVIRLGLMEMIAFKLF----- 242
OY 219 CQLDKDNVQSPGGATIHLFLESGGFRSLINNAVEASCI--RTRELQSMADQEKSIA 276
Db 243 -----CS---GIVSATTEFLESCGQADLT-----TCYGRNKRVAEAPARTGKSIQL 287
OY 277 KKTILDRVKKLESP 289
Db 288 EKEMLNGOKLQGP 300

RESULT 7

US-08-605-150A-6

Sequence 6, Application US/08605150A

Patent No. 6103520

GENERAL INFORMATION:

APPLICANT: Topfer, Reinhard

APPLICANT: Hausmann, Ludger

APPLICANT: Scheil, Jozef

TITLE OF INVENTION: GLYCEROL-3-PHOSPHATE DEHYDROGENASE

TITLE OF INVENTION:

NUMBER OF SEQUENCES: 17

NUMBER OF SEQUENCES: 17

CORRESPONDENCE ADDRESS:

CORRESPONDENCE ADDRESS:

ADRESSEE: Klein & Szekeres

STREET: 4199 Campus Drive, Suite 700

CITY: Irvine

STATE: CA

COUNTRY: USA

ZIP: 92715

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/605,150A

FILING DATE: 01-MAR-1996

CLASSIFICATION: 800

PRIOR APPLICATION DATA:

APPLICATION NUMBER: PCT\EP94\02936

FILING DATE: 02-SEP-1994

PRIOR APPLICATION DATA:

APPLICATION NUMBER: DE P4329827.3
 FILING DATE: 03-SEP-1993
 ATTORNEY/AGENT INFORMATION:
 NAME: Szekeres, Gabor L.
 REGISTRATION NUMBER: 28,675
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 714-854-5502
 FAX: 714-854-4897
 INFORMATION FOR SEQ ID NO: 6:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 367 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-605-150A-6

Query Match 6.3%; Score 98; DB 3; Length 367;
 Best Local Similarity 24.5%; Pred. No. 0.025; Mismatches 121; Indels 44; Gaps 15;
 Matches 70; Conservative 51; Mismatches 121; Indels 44; Gaps 15;

Qy 43 KMGVNLTNSN--KETVKHSDVFLAVKPH---IIPFILDEIGADQVARHIV---VSC 91
 Db 83 KLGRRNVAADPDLLENKDNML--VFTVPHQFVEGICKRKGKIOECAQALSLIKGMEVKM 141
 Qy 92 AAGVTISSEVKKLMAFQAPKVKIRCMNTPVVQEGAT-VVATGTHALVEDGQLE--Q 147
 Db 142 EGPCMISLISDLIG----INCCVNLNGANIANEIAVEKFSEATVGFRENNDIKEKWQ 195
 Qy 148 LMSVSGCTEVEDLIDAVTGLSGSGPAAVMAHD--GGVKGMLPRLAI-OLGAQAL 204
 Db 196 LFSTPYFMVSAVED---VEGVELCGTLKNIVAAAGFVFDGLEMGNTKAIMRIGLREM 251
 Qy 205 LGAAKMLLDSEQ----HPCQQLKD--NVSPGPGATIHALHFLESGGFRSLINAVEASCI 257
 Db 252 KAFSKLLEPSPVQDTFFESCGVADLITLGGRNRKVAEFAKNGNRS--FDDLEAEML 309
 Qy 258 RTRELOQSMADQEKKISPAALKTTLDRVKLE-SPTVSTLTPSSPGKL 302
 Db 310 RGOKLQGVSTAKEV---YEVLRRHGWLEFPLFSTVHEISGRL 350

RESULT 8
 US-08-605-150A-4
 ; Sequence 4, Application US/08605150A
 ; Patent No. 6103520
 ; GENERAL INFORMATION:
 ; APPLICANT: Topfer, Reinhard
 ; APPLICANT: Hausmann, Ludger
 ; APPLICANT: Scheil, Jozef
 ; TITLE OF INVENTION: GLYCEROL-3-PHOSPHATE DEHYDROGENASE
 ; NUMBER OF SEQUENCES: 17
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Klein & Szekeres
 ; STREET: 4199 Campus Drive, Suite 700
 ; CITY: Irvine
 ; STATE: CA
 ; COUNTRY: USA
 ; ZIP: 92715
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patentin Release #1.0, version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/605,150A
 ; FILING DATE: 01-MAR-1996
 ; CLASSIFICATION: 800
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: PCT\BPP94\02936
 ; FILING DATE: 02-SEP-1994

Query Match 6.3%; Score 98; DB 3; Length 381;
 Best Local Similarity 24.5%; Pred. No. 0.026; Mismatches 121; Indels 44; Gaps 15;
 Matches 70; Conservative 51; Mismatches 121; Indels 44; Gaps 15;

Qy 43 KMGVNLTNSN--KETVKHSDVFLAVKPH---IIPFILDEIGADQVARHIV---VSC 91
 Db 97 KLGGNVAADPDLLENKDNML--VFTVPHQFVEGICKRKGKIOECAQALSLIKGMEVKM 155
 Qy 92 AAGVTISSEVKKLMAFQAPKVKIRCMNTPVVQEGAT-VVATGTHALVEDGQLE--Q 147
 Db 156 EGPCMISLISDLIG----INCCVNLNGANIANEIAVEKFSEATVGFRENNDIKEKWQ 209
 Qy 148 LMSVSGCTEVEDLIDAVTGLSGSGPAAVMAHD--GGVKGMLPRLAI-OLGAQAL 204
 Db 210 LFSTPYFMVSAVED---VEGVELCGTLKNIVAAAGFVFDGLEMGNTKAIMRIGLREM 265
 Qy 205 LGAAKMLLDSEQ----HPCQQLKD--NVSPGPGATIHALHFLESGGFRSLINAVEASCI 257
 Db 266 KAFSKLLEPSPVQDTFFESCGVADLITLGGRNRKVAEFAKNGNRS--FDDLEAEML 323
 Qy 258 RTRELOQSMADQEKKISPAALKTTLDRVKLE-SPTVSTLTPSSPGKL 302
 Db 324 RGOKLQGVSTAKEV---YEVLRRHGWLEFPLFSTVHEISGRL 364

RESULT 9
 US-08-375-709-17
 ; Sequence 17, Application US/08375709
 ; Patent No. 5683898
 ; GENERAL INFORMATION:
 ; APPLICANT: YAZAWA, Kazunaga
 ; APPLICANT: TANAKA, Akiko
 ; APPLICANT: KATO, Seishi
 ; APPLICANT: KONDO, Kiyosi
 ; TITLE OF INVENTION: Gene Coding For Eicosapentaenoic Acid
 ; TITLE OF INVENTION: Synthesizing Enzymes and Process for Production of
 ; NUMBER OF SEQUENCES: 19
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Foley & Lardner
 ; STREET: 3000 K Street, N.W., Suite 500
 ; CITY: Washington
 ; STATE: D.C.
 ; COUNTRY: USA
 ; ZIP: 20007-5109
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patentin Release #1.0, version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/375,709
 ; FILING DATE: 20-JAN-1995
 ; CLASSIFICATION: 435

CURRENT FILING DATE: 1998-06-04
 ;
 ; EARLIER APPLICATION NUMBER: 60/048, 650
 ;
 ; NUMBER OF SEQ ID NOS: 66
 ;
 ; SOFTWARE: Patentin Ver. 2.0
 ;
 ; SEQ ID NO: 10
 ;
 ; LENGTH: 543
 ;
 ; ORGANISM: *Shewanella putrefaciens*
 ;
 ; US-09-090-793-10

Query Match 6.2%; Score 97.5; DB 4; Length 543;
 ;
 ; Best Local Similarity 22.5%; Pred. No. 0.053; Mismatches 48;
 ;
 ; Matches 71; Conservative 48; Indels 67; Gaps 13;

Query 8 AGILCAGLIPSRVEAARNRQLALPQGMNLIHSSEPAULRGSEVLEFLKKV 178
 ;
 ; 119 AGILCAGLIPSRVEAARNRQLALPQGMNLIHSSEPAULRGSEVLEFLKKV 178
 ;
 ; Db 54 ETVKRISDVLFLAVKPHITPILDEIGADWQARHIVVS-CAGGTTTSSVEKKLMAFPQAPK 112
 ;
 ; 179 RTVEAS-AFLGQFQIVYRAGLSRDRQGKVVGNKVIAKSVRTEVAEKF- -MPAP- 233
 ;
 ; Db 113 VIRCNTNTPVWQIGATVATGHALVEDQOLIQLQMSVGFCTEVEDLIDAVGLSGS 172
 ;
 ; Db 234 -----AKMLQKLVDDGTTAEMELAQLPVMADDITAEADSG--GH 272
 ;
 ; Db 173 GPAYAFMALDADGGVWKGMLPRLAQLGAQALGAKMLDSQHPCOLKDKNVCPGAT 232
 ;
 ; Db 273 TDNRBLVTL-----LPTILAKKEEQ-----AKYQYDT---PIRVG--CGGSGVT 312
 ;
 ; Db 233 THALHFLESGGFRSLINNAVEASC-----RPRELQSMADQEKISPAALKTLLDRVKL 286
 ;
 ; Db 313 PDAALATFNGAAYITVGCSINQACVEAGASDHTRKLATTEMADVTMAPAADMFMEGVKL 372
 ;
 ; Query 287 ESPVSTLTPSSPGK 302
 ;
 ; Db 373 QVVKRGTLFFMRANKL 388

RESULT 12
 ;
 ; US-08-726-012B-2
 ;
 ; Sequence 2, Application US/08726012B
 ;
 ; Patent No. 5552190
 ;
 ; GENERAL INFORMATION:
 ;
 ; APPLICANT: Hans Joenje, et al.
 ;
 ; TITLE OF INVENTION: CDNA FOR FANCONI ANEMIA COMPLEMENTATION GROUP A
 ;
 ; NUMBER OF SEQUENCES: 18
 ;
 ; CORRESPONDENCE ADDRESS:
 ;
 ; ADDRESSEE: Klarquist Sparkman Campbell Leigh & Winston, LLP
 ;
 ; STREET: One World Trade Center, Suite 1600, 121 S.W. Salmon Street
 ;
 ; CITY: Portland
 ;
 ; STATE: OR
 ;
 ; COUNTRY: USA
 ;
 ; ZIP: 97204-2988
 ;
 ; COMPUTER READABLE FORM:
 ;
 ; MEDIUM TYPE: Disk, 3.5-inch
 ;
 ; COMPUTER: IBM PC compatible
 ;
 ; OPERATING SYSTEM: MS DOS
 ;
 ; SOFTWARE: WordPerfect 5.1+, ASCII
 ;
 ; CURRENT APPLICATION DATA:
 ;
 ; APPLICATION NUMBER: US/08-726, 012B
 ;
 ; FILING DATE: 10/14/96
 ;
 ; CLASSIFICATION: 514
 ;
 ; PRIOR APPLICATION DATA:
 ;
 ; APPLICATION NUMBER:
 ;
 ; FILING DATE:
 ;
 ; ATTORNEY/AGENT INFORMATION:
 ;
 ; NAME: Richard J. Polley
 ;
 ; REGISTRATION NUMBER: 28, 107
 ;
 ; REFERENCE/DOCKET NUMBER: 3812-45520/RJP/DJE
 ;
 ; TELECOMMUNICATION INFORMATION:

Query Match 6.1%; Score 95.5; DB 2; Length 1155;
 ;
 ; Best Local Similarity 22.3%; Pred. No. 0.43; Mismatches 130; Indels 103; Gaps 14;
 ;
 ; Matches 79; Conservative 42; Indels 103; Gaps 14;

Query 19 GILSAHKTIASSPMLVPTVSALRKMGVNLTNSK- 53
 ;
 ; Db 189 GIVSLOBLESHPDMAVGSWLFNRC-CCLCEQMBRASQIADVARMLSDLFVOMFVLRGF 247
 ;
 ; Db 62 -----LELAVKPHIP-----FILDEIGADVO---ARHIVVSCAAGVNTSSE 101
 ;
 ; Db 248 ORNSDLRRTVSEPKMPQVTVDVLQRMLIFALDLAGVQESSTHKIVRCWFGVFSGHTL 307
 ;
 ; Db 102 KKLMAFPAPK-----VIRCNTNTPV-----VQEGATVWVGHALVED----- 141
 ;
 ; Db 308 GSVISTDPKRFPSHLTQIILTHSPVLSKASDAVQMORENSFARTIPLTSLYRLFVMLS 367
 ;
 ; Db 142 -----GOLLEO-----MSSVGFC-TEVEEDLTDAVGLSGSGPAYAFMALD 182
 ;
 ; Db 368 AELVUGHQEVLETOEHWQRLVLSFVSLAVVCFPEAQOLIEDWARM-----AQAFESCO 423
 ;
 ; Db 183 ADGGVKMGLPRLAQLGAQALGAKMLDS-----EOHPCOLKDNV----- 226
 ;
 ; Db 424 LBSMVTATFLVVKQALEGSPASFLSYADWFKASFGSTRGYHCKSKALVFLFTFSELVPE 483
 ;
 ; Db 227 -SPGGTIAHALH-FLESGGFNSLILNAVEASCIRREL---QSMADQERKISPA 274
 ;
 ; Db 484 ESRPYLQVHILHPPLVPSKYRSLLTDYSLAKTRLADLKVISIENMGLYEDLSSA 537

RESULT 13
 ;
 ; US-08-605-150A-8
 ;
 ; Sequence 8, Application US/08605150A
 ;
 ; Patent No. 6103520
 ;
 ; GENERAL INFORMATION:
 ;
 ; APPLICANT: Topfer, Reinhard
 ;
 ; APPLICANT: Hausmann, Ludger
 ;
 ; APPLICANT: Scheil, Jozef
 ;
 ; TITLE OF INVENTION: GLYCEROL-3-PHOSPHATE DEHYDROGENASE
 ;
 ; NUMBER OF INVENTION:
 ;
 ; NUMBER OF SEQUENCES: 17
 ;
 ; CORRESPONDENCE ADDRESS:
 ;
 ; ADDRESSEE: Klein & Szekeres
 ;
 ; STREET: 4199 Campus Drive, Suite 700
 ;
 ; CITY: Irvine
 ;
 ; STATE: CA
 ;
 ; COUNTRY: USA
 ;
 ; ZIP: 92715
 ;
 ; COMPUTER READABLE FORM:
 ;
 ; MEDIUM TYPE: Floppy disk
 ;
 ; COMPUTER: IBM PC compatible
 ;
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ;
 ; SOFTWARE: Patentin Release #1.0, Version #1.25
 ;
 ; CURRENT APPLICATION DATA:
 ;
 ; APPLICATION NUMBER: US/08-605, 150A
 ;
 ; FILING DATE: 01-MAR-1996
 ;
 ; CLASSIFICATION: 800
 ;
 ; PRIOR APPLICATION DATA:
 ;
 ; APPLICATION NUMBER: PCT/EP94\02936
 ;
 ; FILING DATE: 02-SEP-1994
 ;
 ; PRIOR APPLICATION DATA:
 ;
 ; APPLICATION NUMBER: DE P 4329827. 3
 ;
 ; FILING DATE: 03-SEP-1993

TELEPHONE: (503) 226-7391
 ;
 ; TELEFAX: (503) 228-9446
 ;
 ; INFORMATION FOR SEQ ID NO: 2:
 ;
 ; SEQUENCE CHARACTERISTICS:
 ;
 ; LENGTH: 1455
 ;
 ; TYPE: amino acid
 ;
 ; STRANDEDNESS:
 ;
 ; TOPOLOGY: Linear
 ;
 ; US-08-726-012B-2

ATTORNEY/AGENT INFORMATION:
 NAME: Szekeres, Gabor L.
 REGISTRATION NUMBER: 28,675
 REFERENCE/DOCKET NUMBER: 542-04-PA
 TELECOMMUNICATION INFORMATION:
 TELEFAX: 714-854-4897
 INFORMATION FOR SEQ ID NO: 8:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 372 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-605-150A-8

Query Match 6.0%; Score 94.5; DB 3; Length 372;
 Best Local Similarity 23.2%; Pred. No. 0.061; Gaps 18;
 Matches 80; Conservative 52; Mismatches 122; Indels 91; Gaps 18;

QY 23 AKHITASSP-----EMNLPTVSALKRKGVNLTNSN----- 52
 Db 37 AKULIASNLKLPSPFHDEVRMWFEELPSGKLTDV-INQTNENVKYLPGIKLGRNVVA 95
 QY 53 ----KETVKHSDFVLFLAVKPH-----TIPFIDBEGADWQARHIV---VSCAGVNTSS 99
 Db 96 DPDLNEAVKDANML-VFVTPHOFMEGICKRLVGKIOEGAQLSLIKGMEVKMGPOMISS 154
 QY 100 VEVKKLMAFQPAKPVRCMTNTPVVQEGATV-----YATGTHALVEDQOLLE--QL 148
 Db 155 LISDLG-----INC-----CVLMGANIANEIAVEKEFATVFRNTIDIAEKWQL 201
 QY 149 MSSVGFCTEVEEDLIDAVTGLSGSGPAPYFMAKDAD--GGVKGMLPRLAT-OLQAQALL 205
 Db 202 FSTPYFEMSAVED---VEGVELCGTUKNIVIAAGFVDGLEMGNNTKAAMRIGIREMK 257
 QY 206 GAAKMLDSEQ----HPCQDLD-NVCSPPGATIHALHFLFESGGERSLLNAEVASCIR 258
 Db 258 AFSKLFLPSVKDTFFESCGVADLITCLGGRNRKVAAFAKNGGERS-FDDLEABL 315
 QY 259 TRELOSMADOEKS-SPAAIKTLLDRVKL-SPVYSTLTPSSGKL 302
 Db 316 GOKLQGVSTAKEV----YEVIGHRGWLELFPLFSTVHELISTGRL 355

RESULT 14
 US-08-605-150A-10

; Sequence 10, Application US/08605150A
 ; Patent No. 610520

GENERAL INFORMATION:
 APPLICANT: Topfer, Reinhard
 APPLICANT: Hausmann, Ludder
 APPLICANT: Scheil, Jozef
 TITLE OF INVENTION: GLYCEROL-3-PHOSPHATE DEHYDROGENASE
 NUMBER OF SEQUENCES: 17
 TITLE OF INVENTION:
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Klein & Szekeres
 STREET: 4199 Campus Drive, Suite 700
 CITY: Irvine
 STATE: CA
 COUNTRY: USA
 ZIP: 92715

COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patientin Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08-605,150A
 FILING DATE: 01-MAR-1996
 CLASSIFICATION: 800
 PRIOR APPLICATION DATA:

Query Match 6.0%; Score 94.5; DB 3; Length 372;
 Best Local Similarity 23.2%; Pred. No. 0.061; Gaps 18;
 Matches 80; Conservative 52; Mismatches 122; Indels 91; Gaps 18;

QY 23 AKHITASSP-----EMNLPTVSALKRKGVNLTNSN----- 52
 Db 37 AKULIASNLKLPSPFHDEVRMWFEELPSGKLTDV-INQTNENVKYLPGIKLGRNVVA 95
 QY 53 ----KETVKHSDFVLFLAVKPH-----TIPFIDBEGADWQARHIV---VSCAGVNTSS 99
 Db 96 DPDLNEAVKDANML-VFVTPHOFMEGICKRLVGKIOEGAQLSLIKGMEVKMGPOMISS 154
 QY 100 VEVKKLMAFQPAKPVRCMTNTPVVQEGATV-----YATGTHALVEDQOLLE--QL 148
 Db 155 LISDLG-----INC-----CVLMGANIANEIAVEKEFATVFRNTIDIAEKWQL 201
 QY 149 MSSVGFCTEVEEDLIDAVTGLSGSGPAPYFMAKDAD--GGVKGMLPRLAT-OLQAQALL 205
 Db 202 FSTPYFEMSAVED---VEGVELCGTUKNIVIAAGFVDGLEMGNNTKAAMRIGIREMK 257
 QY 206 GAAKMLDSEQ----HPCQDLD-NVCSPPGATIHALHFLFESGGERSLLNAEVASCIR 258
 Db 258 AFSKLFLPSVKDTFFESCGVADLITCLGGRNRKVAAFAKNGGERS-FDDLEABL 315
 QY 259 TRELOSMADOEKS-SPAAIKTLLDRVKL-SPVYSTLTPSSGKL 302
 Db 316 GOKLQGVSTAKEV----YEVIGHRGWLELFPLFSTVHELISTGRL 355

RESULT 15
 US-08-614-377A-7

; Sequence 7, Application US/08614377A
 ; Patent No. 5976864

GENERAL INFORMATION:
 APPLICANT: Smit, John
 APPLICANT: Bingle, Wade H.
 APPLICANT: No. 5976864ellini, John F.
 TITLE OF INVENTION: EXPRESSION AND SECRETION OF
 TITLE OF INVENTION: HETEROLOGOUS
 TITLE OF INVENTION: POLYPEPTIDES FROM CAULOBACTER
 NUMBER OF SEQUENCES: 12
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Fish & Richardson PC
 STREET: 225 Franklin Street
 CITY: Boston
 STATE: Massachusetts
 COUNTRY: USA
 ZIP: 02110-2804

COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS

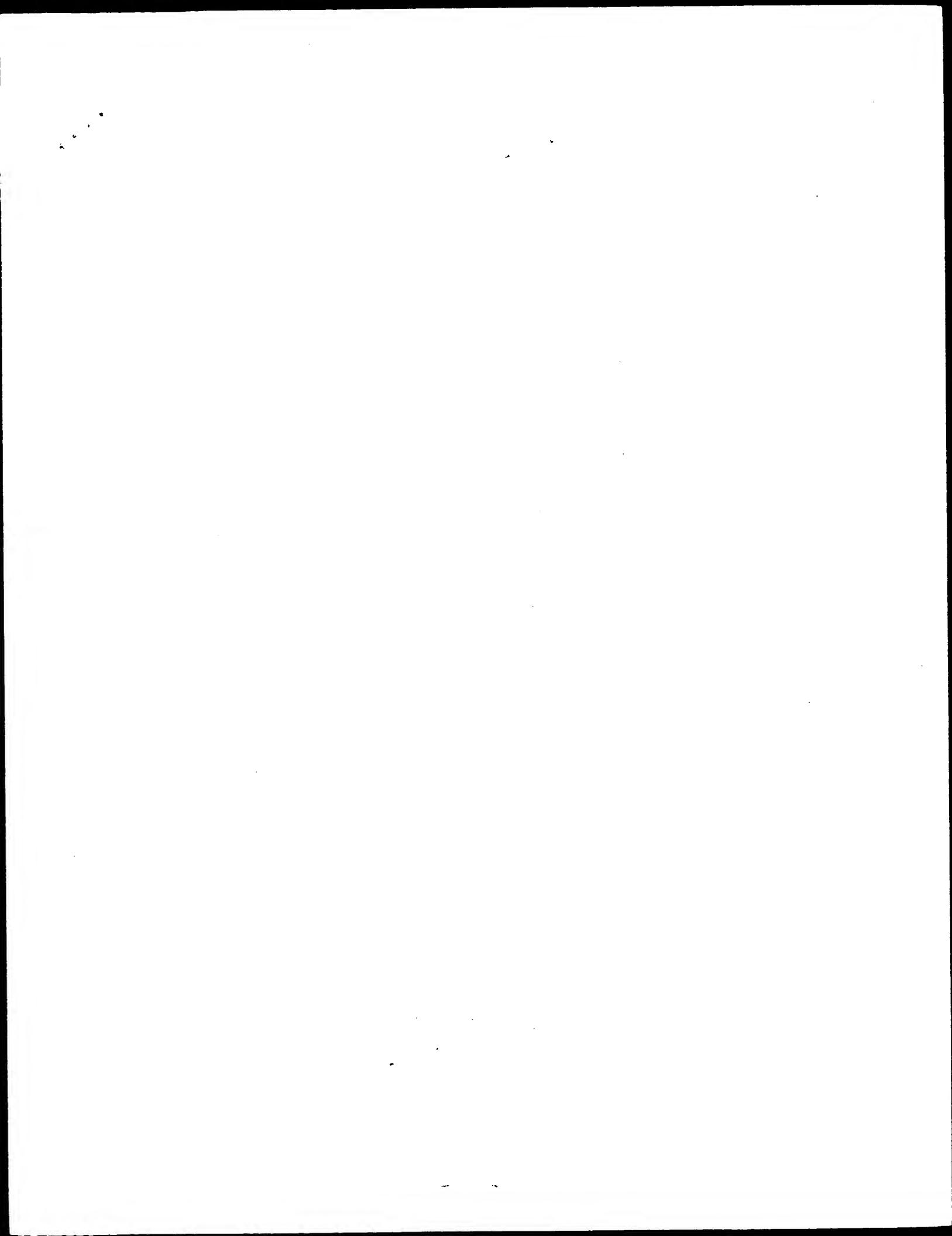
SOFTWARE: PatentIn Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US08/614,377A
 FILING DATE: 12-MAR-1996
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/194,290
 FILING DATE: 09-FEB-1994
 CLASSIFICATION: 435
 APPLICATION NUMBER: US 07/495,367
 FILING DATE: 09-JUNE-1992
 CLASSIFICATION: 435
 ATTORNEY/AGENT INFORMATION:
 NAME: Tsao Y. Rocky
 REGISTRATION NUMBER: 34053
 REFERENCE/DOCKET NUMBER: 01106/002001
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 617-512-5070
 TELEFAX: 617-542-8906
 INFORMATION FOR SEQ ID NO: 7:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1026 amino acids
 TYPE: amino acid
 STRANDEDNESS:
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 ;
 US-08-614-377A-7

Query Match 5 8%; Score 91.5; DB 2; Length 1026;
 Best Local Similarity 22.7%; Pred. No. 67; Mismatches 104; Indels 53; Gaps
 Matches 57; Conservative 37; Mismatches 104; Indels 53; Gaps 10;

```

Qy  27  IASSPENLPTVSALRKMGVLTRSSNKETVKHSVDFLAVKPHITPPFLDEIGADNQARH 86
Db  768  LSSSAALAGCTVALAGVETVNAIATDTNTTAHVDTLQA-----TSAKS 812
Qy  87  IWVSCAAQVITSS-EKKLMAFQ----PAPKVIRCMNTP---VVVOEGATVYATGT 135
Db  813  IWVIGNAGLNLTNTGNTAVTISFDASAVTGATPAVTFVSANTIVGEVVIIRGGA---GA 867
Qy  136  HALVDPGQOLQELMSVGRCTEVEDLIDAVTGS-----GPAVAFMAL-DADGG 186
Db  868  DSLICSATANDTITCAGAGDILVTTGGTDTFTGGTGADIDFDINAGTSAFVITDAAVG 927
Qy  187  VKM---GLPRLAT---QIGAOAIGAARMLLSEQH-CQLKONVCSGGATTHALHLE 240
Db  928  DKLIDLVGISTNGAIDGARGAATL-----AOYLDAAAGDGSSTSVAKWFQ 980
Qy  241  SGGFPSLLINA 251
Db  981  FGGDPTVWVVD 991
  
```

Search completed: January 31, 2002, 14:53:52
 Job time: 22 sec



GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 31, 2002, 14:54:15 ; Search time 12.98 Seconds
(without alignments)
886.961 Million cell updates/sec

Title: US-09-912-717-1

Perfect score: 1567

Sequence: 1 MSVFIGAGQALAYRFTAGI.....TPSSPGKULTRSLALGGKKD 314

Scoring table: BLOSUM62

Gpop 10.0 , Gapext 0.5

Searched: 100059 seqs, 36664827 residues

Total number of hits satisfying chosen parameters: 100059

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0% , Maximum Match 100%

Listing first 45 summaries

Database : Swissprot_39:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description	RESULT	1
1	1306	83.3	319	1	PROC_HUMAN	PROC_HUMAN	STANDARD; PRT; 319 AA.
2	512.5	32.7	276	1	PROC_ARATH	P3322;	01-OCT-1993 (Rel. 27, Created)
3	501.5	32.0	274	1	PROC_SOYBN	004708	01-OCT-1993 (Rel. 27, Last sequence update)
4	479	30.6	273	1	PROC_PEA	004016	20-AUG-2001 (Rel. 40, Last annotation update)
5	476	30.4	278	1	PROC_ARCTH	P74572	DE PYRROLINE-5-CARBOXYLATE REDUCTASE (EC 1.5.1.2) (P5CR) (P5C REDUCTASE).
6	421	26.9	267	1	PROC_SYN3	P00373	GN PYRROLINE-5-CARBOXYLATE REDUCTASE (EC 1.5.1.2) (P5CR) (P5C REDUCTASE).
7	402.5	25.7	269	1	PROC_ECOLT	P22008	OS Homo sapiens (Human).
8	377	24.1	272	1	PROC_PSEAE	P45540	OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo; OX NEBI_TaxID=9606;
9	370	23.6	270	1	PROC_CORG1	J. Biol. Chem. 267:871-875(1992).	RN [1]
10	345	22.0	270	1	PROC_BACSY	RP SEQUENCE FROM N.A.	RR
11	326.5	20.8	294	1	PROC_MICLE	RX MEDLINE=92112821; PubMed=1730675;	RA
12	324.5	20.7	278	1	PROC_BACSU	RT Dougherty K.M., Brandtiss M.C., Valie D.; "Cloning human Pyrroline-5-carboxylate reductase cDNA by complementation in <i>Saccharomyces cerevisiae</i> .";	RT
13	323	20.6	261	1	PROC_THETH	CC -1 - SIMILARITY: BELONGS TO THE PYRROLINE-5-CARBOXYLATE REDUCTASE FAMILY.	RL
14	306.5	19.6	295	1	PROC_MICTU	CC -1 - CATALYTIC ACTIVITY: L-PROLINE + NAD(P)(+) = 1-PYRROLINE-5-CARBOXYLATE + NAD(P)H.	CC
15	301.5	19.2	311	1	PROC_NEUCR	CC -1 - PATHWAY: TERMINAL (THIRD) STEP IN PROLINE BIOSYNTHESIS.	CC
16	297	19.0	265	1	PROC_AQUAE	CC -1 - SUBUNIT: HOMODECAMER OR HOMODECCAMER.	CC
17	296	18.9	320	1	PROC_ZALAR	CC -1 - SIMILARITY: BELONGS TO THE PYRROLINE-5-CARBOXYLATE REDUCTASE FAMILY.	CC
18	281.5	18.0	286	1	PROC_YEAST	CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or_send_an_email_to_licensee@isb-sib.ch).	CC
19	279	17.8	263	1	PROC_TREPA	CC	CC
20	266	17.0	299	1	PROC_GAEEL	CC EMBL; M7835; AAM36407.1; -.	DR
21	264	16.8	271	1	PROC_HAETN	DR PIR; A41770; A41770.	DR
22	252.5	16.1	278	1	PROC_VIBAL	DR MIM: 179035; -.	DR
23	193.5	12.3	257	1	PROC_HELPY	DR InterPro; IPR000304; P5CR.	DR
24	179.5	11.5	257	1	PROC_HELPJ	DR Pfam; PF01099; P5CR; 1.	DR
25	163.5	10.4	273	1	CME4_BACSU	DR PROSITE; PS00521; P5CR; 1.	DR
26	109	7.0	599	1	CENB_MOUSE	DR OXidoreductase; proline biosynthesis, NADP.	DR
27	106	6.8	348	1	GPDA_RABIT	DR SEQUENCE 319 AA; 33374 MW; F5E7B5BDB475DF CRC64;	DR
28	104	6.6	272	1	YKEA_BACSU	DR P5053 vibrio alg1	DR
29	104	6.6	451	1	YVBL_METTP	DR 025773 helicobacte	DR
30	102	6.5	251	1	PROC_METSM	DR 028156 helicobacte	DR
31	102	6.5	348	1	GPDA_MOUSE	DR P22520 methanobrev	DR
32	101	6.4	599	1	CENB_HUMAN	DR P13707 mus musculus	DR
33	6.4	348	1	GPDA_RAT	DR P07199 homo sapien	DR	
					DR 035077 rattus norv	DR	

RESULT 2					
PROC_ARATH		STANDARD:		PRT:	
ID	PROC_ARATH				276 AA.
AC	P54044;				
DT	01-OCT-1996	(Rel. 34, Created)			
DT	01-OCT-1996	(Rel. 34, Last sequence update)			
DT	20-AUG-2001	(Rel. 40, Last annotation update)			
DE	PYRROLINE-5-CARBOXYLATE REDUCTASE (EC 1.5.1.2) (P5CR) (P5C REDUCTASE)				
GN	PROCI OR AT5G14800 OR T9L3-100.				
OS	Arabidopsis thaliana (Mouse-ear cress).				
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eu dicots; core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.				
OC	NCBL_TAXID=3702;				
OX					
RN	[1]				
RN	SEQUENCE FROM N.A.				
RX	MEDLINE-94294559; PubMed=8022935;				
RA	Verbruggen N., Villarreal R., van Montagu M.;				
RA	"Osmoregulation of a Pyrrole-5-carboxylate reductase gene in				
RT	Arabidopsis thaliana";				
RL	Plant Physiol. 103:771-781(1993).				
RN	[2]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN=CV. LANDSBERG ERECTA;				
RC	Villarreal N., Villarreal R., Hua X., van Montagu M.;				
RC	Submitted (OCT-1996) to the EMBL/GenBank/DDBJ databases.				
RC	[3]				
RC	SEQUENCE FROM N.A.				
RA	SEQUENCE FROM N.A.				
RA	MEDLINE-21016721; PubMed=11130714;				
RA	Tabata S., Kaneko T., Nakamura Y., Kotani H., Kato T., Asanizu E.,				
RA	Miyajima N., Sasamoto S., Kimura T., Hosouchi T., Kawashima K.,				
RA	Kohara M., Matsumoto M., Matsuno A., Muraki A., Nakayama S.,				
RA	Nakazaki N., Naruo K., Okumura S., Shinjo S., Takeuchi C., Wada T.,				
RA	Watanabe A., Yamada M., Yasuda M., Sato S., de la Bastide M.,				
RA	Huang E., Spiegel L., Gnoj L., O'Shaughnessy A., Preston R.,				
RA	Habermann K., Murray J., Johnson D., Rohlfing T., Nelson J.,				
RA	Stoneking T., Pepin K., Spielet J., Sekhon M., Armstrong J., Becker M.,				
RA	Bell E., Cordum H., Cordees M., Courtney L., Courtney W., Dante M.,				
RA	Du H., Edwards J., Fryman J., Haakenes B., Lamar E., Latrelle P.,				
RA	Leonard S., Meyer R., Mulvaney E., Ozersky P., Riley A., Strowmatt C.,				
RA	Wagner-Mopherson C., Wollan A., Yoakum M., Bell M., Dedhia N.,				
RA	Parmentier L., Shah R., Rodriguez M., Hoon See L., Vil D., Baker J.,				
RA	Kirchhoff K., Toth K., King L., Baret A., Miller B., Marras M.,				
RA	Martienssen R., McCombie W.R., Wilson R.K., Murphy G., Bancroft I.,				
RA	Volckaert G., Wambutt R., Duesterhoeft A., Stekelen W., Pohl T.,				
RA	Entian K.-D., Terryn N., Hartley N., Bent E., Johnson S., Zimmermann W.,				
RA	Lamhang S.-A., McCullagh B., Robben J., Grysmanprez B.,				
RA	Ransperger U., Wedler H., Balke K., Wedler E., Peters S.,				
RA	van Staveren M., Dirkse W., Moelijman P., Klein Lankhorst R.,				
RA	Weitzelmeier T., Bothe G., Rose M., Hauf J., Bernaeiser S., Hempel S.,				
RA	Volckaert G., Wambutt R., Villarreal R., Gielein J., Ardiles W.,				
RA	Beintz O., Lemcke K., Kolesov G., Mayer K.F.X., Rudd S., Schoof H.,				
RA	Schueler C., Zaccaria P., Mewes H.-W., Bevan M., Fransz P.F.,				
RA	"Sequencing and analysis of chromosomes 5 of the plant Arabidopsis				

RL Nature 408: 823-826(2000).
 CC -!- CATALYTIC ACTIVITY: L-PROLINE + NAD(P)(+) = 1-PYRROLINE-5-
 CC CARBOXYLATE + NAD(P)H.
 CC -!- PATHWAY: TERMINAL (THIRD) STEP IN PROLINE BIOSYNTHESIS.
 CC -!- SUBCELLULAR LOCATION: CYTOPLASMIC.
 CC -!- SIMILARITY: BELONGS TO THE PYRROLINE-5-CARBOXYLATE REDUCTASE
 CC FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>)
 CC or send an email to license@isb-sib.ch).
 CC
 DR EMBL; M76538; CAA61346.1; -.
 DR EMBL; Y08951; CAA70148.1; -.
 DR HSSP; P80028; ITOF.
 DR InterPro; IPR000304; P5CR.
 DR Pfam; PF01089; P5CR; 1.
 DR PROSITE; PS0521; P5CR; 1.
 KW OXIDOREDUCTASE; proline biosynthesis; NADP.
 SQ SEQUENCE 276 AA; 28624 MW; B577A01C92A3A28B CRC64;

 Query Match 32.7%; Score 512.5; DB 1; Length 276;
 Best Local Similarity 42.5%; Pred. No. 4.8e-31; Gaps 4;
 Matches 113; Conservative 52; Mismatches 90; Indels 11; Gaps 4;
 QY 3 VCFIGAQQLAYRF---TAGILSAHKITIASSPEMNLPTVSALKMGVNLTNSRNKEVVKH 58 -
 |||||::|:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:
 Db 13 VCFIGAKMABSIARGVVAAGVGLPPN-ICTAHVNSNINRPROFESFGVNPVTFSEEVKE 71 -
 QY 59 SDVFLFLAVKPHIPIFLDETDGADVQARHIVVSCAAGYTISSEVKLMAFQDPAKVTRCMT 118 -
 ||||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:
 Db 72 SDVVFVFSVKPQVVKAVTELSKLSRNKILVNSVAGKILNDQE---WSGODRFRVMP 127 -
 QY 119 NPPVWVVOEGATVYGHVHALVEDGOLLEFOLMSSVVGCTEVEDDLIDAVTGSGSGPAYA 178 -
 ||||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:
 Db 128 NPPAAYGEAASMSLIGGATTEDGATVAMLFGAVGKLIKADKEMFDATVGLGSGSPAYF 187 -
 QY 179 MALD-ADGGVKGMLPPLRQLQIGAQALIGAKMMLDSEQHPCQQLKDNCSPGGATHAL 236 -
 ||||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:
 Db 188 IIAEALADGGVVAAGLPRLEALASLASQTVLGAATMWSKTGKHPGVLKODVTPGGTTIAGV 247 -
 QY 237 HFLEGGFRRSLLINAVERASCIRTEL 262 -
 ||||:|||:|||:|||:|||:|||:
 Db 248 HELEKGSFRATLMNAVVAAKRSREL 273 -
 ||||:

 RESULT 3
 PROC_SOYBN
 ID SOYBN
 PROC_SOYBN
 STANDARD; PRT; 274 AA.
 AC P17817;
 DT 01-AUG-1990 (Rel. 15, Created)
 DT 01-AUG-1990 (Rel. 15, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE PYRROLINE-5-CARBOXYLATE REDUCTASE (EC 1.5.1.2) (P5CR) (P5C REDUCTASE).
 OS Glycine max (Soybean).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycine.
 OC NCBI_TAXID=3847;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=90340278; PubMed=2199815;
 RA Delaunay A.J., Verma D.P.S.;
 RT "A soybean gene encoding delta 1-pyrroline-5-carboxylate reductase
 was isolated by functional complementation in Escherichia coli and is
 found to be osmoregulated.";

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
CC or send an email to license@lsb-sib.ch).
CC
DR EMBL; U92287; AAC14482; 1; -.
DR InterPro; IPR00304; PSCR.
DR Pfam; PF01089; P5CR; 1.
DR PROSITE; PS00521; P5CR; 1.
DR OXIDOREDUCTASES; Proline biosynthesis; NADP
SEQUENCE 278 AA; 29003 MW; 198B6E34F10236 CRC64;

the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC EMBL; D90916; BAA18679.1; -

CC InterPro; IPR000304; P5CR.

CC Pfam; PF01089; P5CR; 1.

CC PROSITE; PS0521; P5CR; 1.

CC Oxidoreductase; Proline biosynthesis; NADP; Complete proteome.

CC SEQUENCE 267 AA; 27883 MW; FD9C428047F391B0 CRC64;

CC 80

Query	Match	Score	DB	Length
Best Local Similarity	26.9%	421	1	267
Matches	106			
Conservative	38.5%			
Mismatches	47			
Indels	96			
Gaps				
Y	1 MSVGFYGAGQQLA---YRFTAGTLISAHKTIASSPENNLPVVSALR-----KMGVNLTRS	51		
Y	3 IQLGIGGQWMAELLARLIAEKTAPEELIVGEPH-----GARRYDQLQTYQVRVSPD	56		
Y	52 NKETVKHSDVLFYFLAVKPRHIIPIFILDEIGADWQARHIVWSCAGVTTSSVKEKKLMFQAP	111		
Y	57 NQEAAANVEEVILAKVRFQVDRVLSSLAGGAN-RP-IVISITAVGVSQRIORGF---PDH	111		
Y	112 KVIRCMNTNPVWQEGATIVVATGTHALVEDQQL---LEQLMSSVGFCIEVERBDLAVTGL	169		
Y	112 ALIRAMPNTPAV--GAGMTAAANKWVPEQDOLAKAKAIFRSVGNVNEPVNPLMDAVTG	169		
Y	170 SGSGPRAFYAFMALLD--ADGGYKMGFLPRLATQIGAALIGAKMFLDSEQHCKLQKNC	227		
Y	170 SSGSGPAPAYVALMIEALADGGVYAGLPLPRAIAORQALQTVLGLTAELIKEETEHPAQIKDVT	227		
Y	228 PGGTAIHALHFFLESQGFRSLILINAVEASCCRTEL	262		
Y	230 FGTTTIAVQWKGKLFQFSRSLTEAVRAAYRERSOEL	264		

RA Lin D., Namath A., Oefner P., Roberts D., Schramm S., Davis R.W.;
 RL Submitted (NOV-1995) to the EMBL/GenBank/DBJ databases.
 CC -I CATALYTIC ACTIVITY: L-PROLINE + NAD(P)(+)= 1-PYRROLINE-5-
 CC CARBOXYLATE + NAD(P)H.
 CC -I PATHWAY: TERMINAL (THIRD) STEP IN PROLINE BIOSYNTHESIS.
 CC -I SIMILARITY: BELONGS TO THE PYRROLINE-5-CARBOXYLATE REDUCTASE
 FAMILY.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL: J01605; AAA86433.1; -.
 DR EMBL; AA000145; AAC73489.1; -.
 DR EMBL; U73857; RAB18110.1; -.
 DR HSSP; P14295; 1HYH.
 DR Ecogene; EG10765; PROC.
 DR InterPro; IPR000304; P5CR.
 DR Pfam; PF0089; P5CR; 1.
 DR PROSITE; PS00521; P5CR; 1.
 DR KW Oxidoreductase; Proline biosynthesis; NADP; Complete proteome.
 SO SEQUENCE 269 AA; 28145 MW; C5E2CED36BCC508D CRC64;

Query Match 25.7%; Score 402.5; DB 1; Length 269;
 Best Local Similarity 35.5%; Pred. No. 6.2e-23; Indels 19; Gaps 6;
 Matches 98; Conservative 52; Mismatches 107; Indels 19; Gaps 6;

Qy 3 VGFIGAGOLAYRFTAGILSAHKITASSPEMNLPT--VSALR-KMGVLNTRSNKRNKTVKH 58
 Dr 5 IGFICGNGM-KAIGGLIAGSOVINGQWVYVTPSPDKVAAHLHQFGINAEQAEVQI 63

Qy 59 SDVLFFLAKPHIPHTLDEIGADVOARHTVSCAGVTISVERKIMAFQAPAKVIRCKM 118
 Dr 64 ADIFRAVKPGIMKVLSETSSLNKSLSVIAQVTLQDLR--ALGHDRKTRAMP 120

Qy 119 NTPVVVQEGAVVYAGTHALVEDGQLEQMLSSVQGFCTEVEEDLIDAVGVLSSGGPAAF 178
 Dr 121 NTPPALVNVAGMPTSVTINALVTPEDATDVLNTRFCFGAEAVIAEPMTHPVVGVGSSSSV 180

Qy 179 MALD--ADGGVKGMPRRRLAIQLGQALLGAKMDSQHPCOLKDNYCSPGGATIH 236
 Dr 181 MFIEAMADAVALGGMPRAQAYKFAAQAVNGSAKMYLDEGEHGPGLKDMCSPGGTIEAV 240

Qy 237 HFLESGGFRSLINNAVEASCIRTRELQSMADQEKT S 272
 Dr 241 RYLERKGFRAVIEAM-----TKCMEKSEKL S 267

RESULT 8

PROC_PSEAE ID PROC_PSEAE STANDARD PRT 272 AA.

DT P22008; 01-AUG-1991 (Rel. 19, Created)
 DT 01-AUG-1991 (Rel. 19, Last sequence update)
 DT 20-AUG-2001 (Rel. 40, Last annotation update)
 DE PYRROLINE-5-CARBOXYLATE REDUCTASE (EC 1.5.1.2) (P5CR) (P5C REDUCTASE).
 GN PROD OR P0393.
 OS Pseudomonas aeruginosa.
 OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
 OC Pseudomonas.
 OC NCBI_TaxID=287;
 RN [1]
 RP SEQUENCE FROM N.A., AND SEQUENCE OF 1-12.
 RC STRAIN=ATCC 15592 / PAO1;
 RX MEDLINE=90185238; PubMed=2107123;
 RA Savioz A., Jeenes D.J., Kocher H.P., Haas D.;
 RA "Comparison of proc and other housekeeping genes of Pseudomonas

RT aeruginosa with their counterparts in Escherichia coli.",
 RL Gene 86:107-111(1990).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 15692 / PAO1;
 RX MEDLINE=20437337; PubMed=11984043;
 RA Stover C.K., Pham X.-Q.T., Erwin A.L., Mizoguchi S.D., Warrener P.,
 RA Hickey M.J., Brinkman F.S.L., Hufnagle W.O., Kowalk D.J.J., Lagrou M.,
 RA Gardner R.L., Gohtry L., Tolentino E., Westbroek-Wadman S., Yuan Y.,
 RA Brody L.L., Coulter S.N., Folger K.R., Kas A., Larbig K., Lin R.M.,
 RA Smith K.A., Spencer D.H., Wong G.K.S., Wu Z., Pausen I.T.,
 RA Reizer J., Sauer M.H., Hancock R.E.W., Lory S., Olson M.V.;
 RT "Complete genome sequence of *Pseudomonas aeruginosa* PAO1, an
 RT opportunistic pathogen.",
 RL Nature 406:959-964(2000).
 RN [3]
 RP SEQUENCE OF 1-6 FROM N.A.
 RC STRAIN=ATCC 15692 / PAO1;
 RX MEDLINE=91285432; PubMed=1576385;
 RA Whitechurch C.B., Hobbs M., Livingston S.P., Krishnapillai V.,
 RA Mattick J.S.;
 RT "Characterisation of a *Pseudomonas aeruginosa* twitching motility gene
 RT and evidence for a specialised protein export system widespread in
 RT eubacteria.",
 RL Gene 101:33-44(1991).
 RC -I CATALYTIC ACTIVITY: L-PROLINE + NAD(P)(+)= 1-PYRROLINE-5-
 CC CARBOXYLATE + NAD(P)H.
 CC -I PATHWAY: TERMINAL (THIRD) STEP IN PROLINE BIOSYNTHESIS.
 CC -I SIMILARITY: BELONGS TO THE PYRROLINE-5-CARBOXYLATE REDUCTASE
 CC FAMILY.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC DR EMBL; M33537; AAA25975.1; -.
 DR EMBL; AA004476; AA03782.1; -.
 DR EMBL; M5524; AA25958.1; -.
 DR PIR; JQ0418; JQ0418.
 DR InterPro; IPR000304; P5CR.
 DR Pfam; PF01089; P5CR; 1.
 DR PROSITE; PS00521; P5CR; 1.
 DR KW Oxidoreductase; Proline biosynthesis; NADP; Complete proteome.
 FT INIT_MER 0
 FT SEQUENCE 272 AA; 27963 MW; 750520E6028D73E0 CRC64;

Query Match 24.1%; Score 377; DB 1; Length 272;
 Best Local Similarity 36.6%; Pred. No. 4.8e-21; Indels 18; Gaps 10;
 Matches 101; Conservative 53; Mismatches 104; Indels 18; Gaps 10;

Qy 3 VGFIGAGOLAYR---PTAAGILSAHKITASSPEMNLPTVSALRKMGVNLTTRSNKRNKTVKH 58
 Dr 5 IAFIGAGNMASLIGGLIQRAGVPA-QIRASDPGE-QRAKIAGEFAIDVVEVSNAEAVAD 62

Qy 59 SDVLFFLAKPHIPHTLDEIGADVOARHTVSCAGVTISVERKIMAFQAPAKVIRCKM 118
 Dr 63 ADYVFLVLSVKPQAMKAVQCALAPALKPQLVITGIPCAASLWLL--GQPRP-VVRCM 119

Qy 119 NTPVVVQEGAT-WIAGTHALVEDGQLE-QLMSSVGFCTEVEEDL-IDAWTGVLSSGP 174
 Dr 120 NTPFALLRQGASOLYA---NAQVSAADGEOAQGOLLSAUGIALWLDDEAQIDAVTAVSGSP 176

Qy 175 AVFALMAD--ADGGVKGMPRLAIQLGQALLGAKMDSQHPCOLKDNYCSPGGAT 232
 Dr 177 AYFLLQAMTPTAGEKIGLSRSRASRILQTLQALGAQMAMLSSEVEPAELRRRVSPNGT 236

Qy 233 IHALHFLESGGGRSLINNAVEASCIRTRELQSMAD 268

Db 237 EAIKSFOANGFEALVEQALNAASORSALAEOLGQ 272

RESULT 9

PROC_CORGL STANDARD; PRT; 270 AA.

ID P46540; AC P46540; DT 01-NOV-1995 (Rel. 32, Created)

DT 01-NOV-1995 (Rel. 32, Last sequence update)

DT 20-AUG-2001 (Rel. 40, Last annotation update)

DE PYRROLINE-5-CARBOXYLATE REDUCTASE (EC 1.5.1.2) (P5PCR) (P5C REDUCTASE).

GN PROC.

GS Corynebacterium glutamicum (Brevibacterium flavidum).

OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;

OC Actinomycetales; Corynebacterineae; Corynebacteriaceae;

OC Corynebacterium.

OX NCBI_TaxID=1718;

RN [1] SEQUENCE FROM N.A.

RC STRAIN=ATCC 17985 / Melassecola;

RX MEDLINE=96345604; PubMed=8755067;

RA Ankri S., Serebrjiski I., Reyes O., Leblon G.;

RT "Mutations in the Corynebacterium glutamicum proline biosynthetic pathway: a natural bypass of the proA step.", J. Bacteriol. 178: 4412-4419 (1996).

-!- CATALYTIC ACTIVITY: L-PROLINE + NAD(P)(+)= 1-PYRROLINE-5-CARBOXYLATE + NAD(P)H

CC -!- PATHWAY: TERMINAL (THIRD) STEP IN PROLINE BIOSYNTHESIS.

CC -!- SIMILARITY: BELONGS TO THE PYRROLINE-5-CARBOXYLATE REDUCTASE FAMILY.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC -----

CC EMBL; U31225; AAC44172.1; -.

DR InterPro; IPR000304; P5CR.

DR Pfam; PF01089; P5CR; 1.

DR PROSITE; PS00521; P5CR; 1.

DR OXIDOREDUCTASE; Proline biosynthesis; NADP.

KW Sequence 270 AA; 28223 MW; 90150E233D94158F CRC64;

Query Match 23.6%; Score 370; DB 1; Length 270;

Best Local Similarity 33.8%; Pred. No. 1 6e-20; DB 1; Length 270;

Matches 93; Conservative 56; Mismatches 102; Indels 24; Gaps 8;

Qy 2 SVGFIGAGQOLAYRFTAGLISAHKLTASSPENLPTVSALRKMG-----VNLTNS 52

Db 3 TAVIGGQGIG-EALVSGLIAN---MNPQ-NIRVTRNSBERGGLDRYGLIIMT-DN 55

Qy 53 KETVKHSVLFIAVKPHIPIFILDE1GADV--QARHIVVSCAAGTISVEKKLMAFQ 109

Db 56 SQAADAEDEVFLCVKPKEFIVELTSEITGTLDNNSAQSVVVMSAAGISIAAMEASAGPL 115

Qy 110 ARKVIREMTNTPVVWEGATVATGTHALVEDQGLEQIQLMSSVNGFCTEVEDLIDAVTG 169

Db 116 --WVRIMPNTPLVKGMSVTKGKVDAEQLEQWQKDLSTGVQLEVAESDIDAVTM 172

Qy 170 SGGPAPYAFMADA--DGGVKGGLPRLATQDGAQALGAAKMLDSEOHPCOLKDVC 227

Db 173 SSSSPAYFLFLVATEALITAGVNDGLPRTAKKAVASFEGATMIMKETGKPESELPRAGV 232

PROH_BACSU

ID PROH_BACSU STANDARD; PRT; 270 AA.

AC P14383; 031828;

DT 01-JAN-1990 (Rel. 13, Created)

DT 15-JUL-1998 (Rel. 36, Last sequence update)

DT 20-AUG-2001 (Rel. 40, Last annotation update)

DE PYRROLINE-5-CARBOXYLATE REDUCTASE HOMOLOG 1.

GN YOXE OR PROC.

OS Bacillus subtilis.

OC Bacteria; Firmicutes; Bacillus/Clostridium group;

OC Bacillus/Staphylococcus group; Bacillus.

OX NCBI_TaxID=1423;

RN [1] SEQUENCE FROM N.A.

RP SEQUENCE FROM N.A.

RC STRAIN=168;

RX MEDLINE=91196601; PubMed=1849493;

RA Ahn K.S., Wake R.G.;

RT "Variations and coding features of the sequence spanning the replication terminus of *Bacillus subtilis* 168 and W23", Gene 98:107-112(1991).

RL [3] SEQUENCE OF 1-200 FROM N.A.

RX MEDLINE=89155440; PubMed=2493444;

RA Lewis P.J., Wake R.G.;

RT "DNA and protein sequence conservation at the replication terminus in *Bacillus subtilis* 168 and W23.", J. Bacteriol. 171:1402-1408(1989).

CC -!- FUNCTION: NOT KNOWN, BUT COULD BE INVOLVED IN GLUTAMINE BIOSYNTHESIS.

CC -!- SIMILARITY: BELONGS TO THE PYRROLINE-5-CARBOXYLATE REDUCTASE FAMILY.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC -----

DR EMBL; Z9914; CAB13741.1; -.

DR EMBL; M24523; AAB22722.1; -.

DR PIR; B32807; B32807.

DR Subtilist; BG1109; YOXE.

DR InterPro; IPR000304; P5CR.

DR Pfam; PF01089; P5CR; 1.

DR PROSITE; PS00521; P5CR; 1.

KW Hypothetical protein; OXIDOREDUCTASE; complete proteome.

FT VARIANT 58 58 T->A (IN STRAIN W23).

FT VARIANT 70 70 L->S (IN STRAIN W23).

FT VARIANT 74 74 L->I (IN STRAIN W23).

FT VARIANT 93 93 N->S (IN STRAIN W23).

FT VARIANT 100 100 S->T (IN STRAIN W23).

FT VARIANT 127 127 E->Q (IN STRAIN W23).

FT VARIANT 159 159 K->Q (IN STRAIN W23).

FT VARIANT 236 236 H->Q (IN STRAIN W23).

FT VARIANT 239 239 I->V (IN STRAIN W23).

FT CONFLICT 256 256 Q->E (IN REF. 2).

SQ SEQUENCE 270 AA; 29402 MW; C7B9E1680B7FF9A7 CRC64;

Query Match 22.0%; Score 345; DB 1; Length 270;

Best Local Similarity 34.6%; Pred. No. 1 1e-18;

Matches 88; Conservative 53; Mismatches 101; Indels 12; Gaps 6;

Qy 3 VGF1GAGQOLAYRFTAGLISAHKLTASSPENLPTVSALRKMG-KGVNLTNSNKIV 57

Db 233 PAGTVAIRELEESGIRGAFYRAQACDRSEEL 267

RESULT 10

Db 19' VAFIGAGSMA-BGMMISGLVRANKIPKONICVNRNSWPERTELEQIGKALPHQLCIE 77
 Qy 58 HSDVLFIAVKPHIPIFILDEIAGDVAHRTIVSCAAGVTISVKEKMAOPAPKVRCM 117
 Db 78 DKDVLLILAKPKDADENALSSKSRIOPHOLLSVLAGITSFIEQSLNEQP--VVRLW 134
 Qy 118 TWPVVVYQEGAVVYAVGTHALVEDGOLLEQMSVGFCTEVFEDIDAVGLSGSPAYA 177
 Db 135 PTWTSMIGASATAILALGKVSYEDLKLAEALLGICGEVYTIQENQMDIFGIAAGSPAYF 194
 Qy 178 FMAALD--ADGGVKGMLPRLAQLQAGAKMLDSEOPCOKLNDVNCSPGGATHA 235
 Db 195 YVLMERIEKTEGEEGLDKQUSRSIGAQTLGAAKLMETGEHPELIRDNTSPNQTTAG 254
 Qy 236 LHFL-ESGGFRSL 248
 Db 255 LQALKNSGGERFL 268

RESULT 11
 PROC_MVCLE STANDARD; PRT; 294 AA.
 ID P46725:
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 20-AUG-2001 (Rel. 40, Last annotation update)
 DE PYRROLINE-5-CARBOXYLATE REDUCTASE (EC 1.5.1.2) (P5CR) (P5C REDUCTASE).
 GN PROC OR ML2430 OR B2168_C2_211.
 OS Mycobacterium leprae.
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
 OC Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.
 OX NCBI_TaxID=1769;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Smith D.R.; Robison K.;
 RL Submitted (MAR-1994) to the EMBL/GenBank/DDBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=TN;
 RX MEDLINE=21128732; PubMed=11234002;
 RA Cole S.T., Elgmeier K., Parkhill J., James K.D., Thomson N.R.,
 RA Wheeler P.R., Honore N., Garner T., Churcher C., Harris D.,
 RA Mungall J., Basham D., Brown D., Chillingworth T., Connor R.,
 RA Davies R.M., Devlin K., Duthoy S., Feltwell T., Fraser A., Hamlin N.,
 RA Holroyd S., Hornsby T., Jageis K., Lacroix C., Maclean J., Moule S.,
 RA Murphy L., Oliver K., Quail M.A., Rajandream M.A., Rutherford K.M.,
 RA Rutter S., Seeger K., Simon S., Simmonds M., Skelton J., Squares R.,
 RA Squares S., Stevens K., Taylor K., Whitehead S., Woodward J.R.,
 RA Barrell B.G.;
 RT "Massive gene decay in the leprosy bacillus.";
 RL Nature 409:1007-1011(2001).
 CC -1- CATALYTIC ACTIVITY: L-PROLINE + NAD(P)(+)= 1-PYRROLINE-5-
 CC CARBOXYLATE + NAD(P)H.
 CC -1- PATHWAY: TERMINAL (THIRD) STEP IN PROLINE BIOSYNTHESIS.
 CC -1- SIMILARITY: BELONGS TO THE PYRROLINE-5-CARBOXYLATE REDUCTASE
 CC FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 DR EMBL; ALB83925; CAC31947; 1; -.
 DR Leproma; ML2430; -.
 DR InterPro; IPR000304; P5CR.
 DR Pfam; PF01089; P5CR; 1.
 DR PROSITE; PS00521; P5CR; 1.
 KW Oxidoreductase; Proline biosynthesis; NADP; Complete proteome.
 SEQUENCE 294 AA; 30237 MW; E860659CBB6D9D CRC64;

Query Match 20.8%; Score 326.5; DB 1; Length 294;
 Best Local Similarity 36.2%; Pred. No. 2.8e-17; Gaps 7;
 Matches 88; Conservative 54; Mismatches 113; Indels 39; Gaps 7;
 DR 48 LTRSNKETVKHSDVLFIAVKPHIPIFILDEI---GADVQAHIVVSCAAGVTISVEK 102
 DR 54 LVTSYTDAVENAMFVWVAVKPTQEVESVAGDLQAAVANDSAEQVLTVAAGTIVLES 113
 Qy 103 KLMQFOPAPKVTRGNTPVVWQEGATWAT3THALVEDGOLLEQMS--SVGFCTEVF 159
 Db 114 KLPACTP--VVRAMPNAALVAGVAGTVLAKGRV--TGQNFEDVLMFDIAVGGVLT 167
 Qy 160 EDLIDAVGLGSGSPAYFEMALDA--DGGVKGMLPRLAQLQAGAQLGARNML--DS 214
 Db 168 ESQMDAVTAVSGSGSPAYFELLVLEALVDAULQROVATELAMQTMAGSAMLLRMDQ 227
 Qy 215 EQH-----PCQLKDNCSPGGATHALFLESGCRFRSLINAVNEASCIRTR 260
 Db 228 DRHSAEVAPLGAQDVPAQLRATITSPGTTAAALRERGLRMVDAAVQAKIRSE 287
 Qy 261 BLQ 263
 DR 288 QLR 290

RESULT 12
 PROBACSU STANDARD; PRT; 278 AA.
 ID PROBACSU
 AC P54552;
 DT 01-OCT-1995 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 20-AUG-2001 (Rel. 40, Last annotation update)
 DE PYRROLINE-5-CARBOXYLATE REDUCTASE HOMOLOG 2.
 GN YQ00
 OS Bacillus subtilis.
 OC Bacteria; Firmicutes; Bacillus/clostridium group;
 OC Bacillus/Staphylococcus group; Bacillus.
 RN NCBI_TaxID=1423;
 RC STRAIN=168 / JH642;
 RA Koyashii Y., Micuno M., Masuda S., Takemaru K., Hosono S.,
 RA Sato T., Takeuchi M.;
 RL Submitted (MAY-1996) to the EMBL/GenBank/DDBJ databases.
 CC -1- SIMILARITY: BELONGS TO THE PYRROLINE-5-CARBOXYLATE REDUCTASE
 CC FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR D84432; BAM12621; 1; -.
 DR EMBL; 299116; CAC14312; 1; -.
 DR Subtilist; BG11744; yqj0.
 DR InterPro; IPR000304; P5CR.
 DR Pfam; PF01089; P5CR; 1.
 DR PROSITE; PS00521; P5CR; 1.
 KW Hypothetical protein; Oxidoreductase; Complete proteome.
 SEQUENCE 278 AA; 30396 MW; 4B69DB527E55345E CRC64;

Query Match 20.7%; Score 324.5; DB 1; Length 278;
 Best Local Similarity 29.9%; Pred. No. 3.7e-17; Gaps 7;
 Matches 88; Conservative 54; Mismatches 113; Indels 39; Gaps 7;

Qy 3 VAFIGAGOLAYR---FTAAGILSAKHTIASPEENLPTVSLARKMGVNUTRSNKTVKH 58

DR Pfam; PF01089; P5CR; 1.
 DR PROSITE; PS0521; P5CR; 1.
 KW Oxidoreductase; Proline biosynthesis; NADP; Complete proteome.
 SEQUENCE 295 AA; 30171 MW; F95C3F407BE5408F CRC64;

Query Match 19.6%; Score 306.5; DB 1; Length 295;
 Best Local Similarity 34.3%; Pred. No. 8.5e-16; Matches 86; Conservative 39; Mismatches 91; Indels 35; Gaps 6; Matches 81; Conservative 49; Mismatches 102; Indels 23; Gaps 6;

Qy 48 LTRSNKETVKVSKDVLFLAVKP-HIIPFILDEIGADV---QARRHIVVSCAACVTSVSE 101
 Db 54 LVTSAADAVENATEFVVAVKPAPDVBPVIALANATAAENDSAEQFVTVWAGITVYFE 113
 Qy 102 KKLMAFQAPAKVIRMTNTPVVQDQCATYATGTHALVEDGQOLQMLSSVQFCCTEVEED 161
 Db 114 SKLPACTP---WVAMPAAALVAGVATLAKRFRVTPQQLERSVALEDAVGGVLTYPES 170
 Qy 162 LIDAVTGLSGSGPQYAFMALDA-DGGVWGLRPLATOLGQALIGAAKMLDSOH-- 217
 Db 171 OLDAVTAVSGSGPQAVFLFLVLEALDAGVQVGLRSRQVATLQNTMAGSAMMLERMEQD 230
 Qy 218 -----PCQLDNYCSPGGATTHALFLESGGFRSLINAVASCIRTEL 262
 Db 231 GGANGELMLGRVDTIASRIRAVTSPGQTAAALRELERGFRMAYDAVAQA----- 283
 Qy 263 QSMADQEKITSP 273
 Db 284 KSRSQLRITP 294

RESULT 15

PROC_NEUCR	STANDARD;	PRT;	311 AA.
ID 012641;			
AC 012641;			
DT 15-JUL-1998 (Rel. 36, Created)			
DT 15-JUL-1998 (Rel. 36, Last sequence update)			
DE PYRROLINE-5-CARBOXYLATE REDUCTASE (EC 1.5.1.2) (P5CR) (P5C REDUCTASE).			
GN PRO-1.			
OS Neurospora crassa.			
OC Sordariidales; Sordariaceae; Neurospora.			
OX NCBI_TaxID:5141;			
RN [1]			
RP SEQUENCE FROM N.A.			
RX MEDLINE=96004701; PUBMED=2565396;			
RA Davis C. R., McPeek M. A., McClung C. R.,			
RT "Molecular characterization of the proline-1 (pro-1) locus of			
RT Neurospora crassa, which encodes delta 1-pyrroline-5-carboxylate			
RT reductase.",			
RL Mol. Gen. Genet. 248:341-350(1995).			
CC -!- CATALYTIC ACTIVITY: L-PROLINE + NAD(P)(+)= 1-PYRROLINE-5-CARBOXYLATE + NAD(P)H.			
CC -!- PATHWAY: TERMINAL (THIRD) STEP IN PROLINE BIOSYNTHESIS.			
CC -!- SIMILARITY: BELONGS TO THE PYRROLINE-5-CARBOXYLATE REDUCTASE FAMILY.			
CC			

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

DR InterPro: IPR000304; P5CR.

DR Pfam: PF01089; P5CR; 1.

DR PROSITE; PS0521; P5CR; 1.

KW Oxidoreductase; Proline biosynthesis; NADP.

SQ SEQUENCE 311 AA; 32151 MW; 9A6D525D845984DF CRC64;

Search completed: January 31, 2002, 14:57:38
 Job time: 203 sec

Thu Jan 31 15:07:58 2002

us-09-912-717-1.rsp

Copyright (c) 1993 - 2000 Compugen Ltd.	GenCore version 4.5			
Run on:	January 31, 2002, 14:53:30 ; Search time 16.15 Seconds			
Title: Perfect score:	US-09-912-717-1			
Sequence:	1 MSVFIGAGQALYRFTAAGI TPSSPGKLLTRSLALGGKD 314			
Scoring table:	BLOSUM62			
Gappen:	10.0 , Gapext 0.5			
Searched:	219241 seqs, 76174552 residues			
Total number of hits satisfying chosen parameters:	219241			
Minimum DB seq length:	0			
Maximum DB seq length:	200000000			
Post-processing:	Minimum Match 0%			
	Maximum Match 100%			
	Listing first 45 summaries			
Database :	PIR_68:*			
	1: PIR1:*			
	2: PIR2:*			
	3: PIR3:*			
	4: PIR4:*			
Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.				
	SUMMARIES			
Result No.	Query	Length	DB ID	Description
1	1306	83.3	319	1 A41770
2	578	36.9	279	2 T23765
3	52.5	32.7	276	1 JQ2334
4	501.5	32.0	274	1 S10186
5	479	30.6	273	2 T06477
6	423	27.0	284	1 T36286
7	26.9	26.7	272	1 T6767
8	405.5	25.9	269	2 H85533
9	402.5	25.7	269	1 R0BCC
10	377	24.1	273	2 J00418
11	345	22.0	270	2 E59682
12	326.5	20.8	274	1 S72897
13	344.5	20.7	278	1 G69964
14	323.5	20.6	261	2 JQ2078
15	323.5	20.6	279	2 G83837
16	319	20.4	255	2 E72360
17	306.5	19.6	295	2 G70745
18	301.5	19.2	311	2 S57863
19	297.5	19.0	305	2 C82524
20	297	19.0	265	2 F70315
21	296	18.9	320	2 JC4830
22	295.5	18.9	262	2 D86860
23	293	18.7	266	2 G72769
24	292.5	18.7	282	2 P50305
25	289	18.4	264	2 C75385
26	281.5	18.0	286	2 S52593
27	279	17.8	263	2 D17281
28	266	17.0	299	2 T29226
29	264	16.8	271	2 I64060

RESULT 2

T23765 hypothetical protein M153.1 - *Caenorhabditis elegans*

C;Species: *Caenorhabditis elegans*

C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 18-Feb-2000

C;Accession: T23765

R;Matthews, P.

A;Reference number: 219796

A;Introns: 67/3; 84/1; 114/3; 139/1; 179/2; 214/3

A;Note: T9L3.100

A;Status: Preliminary; translated from GB/EMBL/DDJB

A;Molecule type: DNA

A;Residues: 1-279 <WIL>

A;Cross-references: EMBL:Z67995; PIDN:CAA91943.1; GSPDB:GN00028; CESP:M153.1

A;Experimental source: clone M153

C;Genetics:

A;Gene: CESP:M153.1

A;Map position: X

A;Introns: 23/1; 106/3; 177/3; 248/3

C;Superfamily: Pyrroline-5-carboxylate reductase

Query Match 36.9%; Score 578; DB 2; Length 279;

Best Local Similarity 44.6%; Pred. No. 1.3e-38;

Matches 127; Conservative 55; Mismatches 89; Indels 14; Gaps 4;

QY 1 MSVFIGAGOLAYRFTA---AGIUSAHKTIASSEMNLPTVSALRKMGVNLTNSKETV 56

Db 1 MKIGFIGAGKMAQIARGLINSGRTRADNTIASSPKRDEFVLDQKALGINTHNAEV 60

Db 61 QKSDVFLFLAVKPVHVKVASEIAPALKERHLYWNLSTALGITIRNIESLL--PTKSRVVR 117

QY 117 MNTPVWVQEGATVYVGTGHALVEGQLELQMLSSVGCFTEVEEDLIDAVTGLGSGPAPV 176

Db 118 MNTPVPSVVRAGAFAMGSACRDGAETVERKLSTVGFPEVPEHIDPVTGSGSPY 177

QY 177 AFMFLAD-ADGGVKMGLPRLAQLQGAQALGAQMLDSEHQPCQQLDKWCSPGGATH 234

Db 178 MPAVFLSLADGGVKMGLPRLAQLKAAATYPLGAKRNLHQPLQDQVSPGGSY 237

QY 235 ALHFLESGGGFRLSLINAVASCIRTELQESMADEKISPAKKT 279

Db 238 GHMKLLEGGGLKVLMDAVEAATNRSR----ATGDKALPDRFRNT 277

RESULT 3

JQ2334 pyrroline-5-carboxylate reductase (EC 1.5.1.2) [similarity] - *Arabidopsis thaliana*

C;Species: *Arabidopsis thaliana* (mouse-ear cress)

C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999

C;Accession: JQ2334; T51425

R;Verbruggen, N.; Vilarruel, R.; Van Montagu, M.

A;Title: Osmoregulation of a pyrroline-5-carboxylate reductase gene in *Arabidopsis thaliana*

A;Reference number: JQ2334; MUID:94294559

A;Accession: JQ2334

A;Molecule type: DNA

A;Residues: 1-276 <VER>

A;Cross-references: GB:MR6538; NID:9166814; PIDN:AAA61346.1; PID:9166815

R;Sato, S.; Nakamura, Y.; Kaneko, T.; Kato, T.; Asamizu, E.; Kotani, H.; Tabata, S.; Mew

A;Submitted to the Protein Sequence Database, August 2000

A;Reference number: 225394

A;Accession: T51425

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-276 <SAM>

A;Cross-references: EMBL:AL391149

A;Experimental source: cultivar Columbia; BAC clone T9L3

C;Comment: This enzyme is involved in the last step of the proline biosynthetic pathway.

C;Genetics:

A;Map position: 5

A;Introns: 67/3; 84/1; 114/3; 139/1; 179/2; 214/3

A;Note: T9L3.100

A;Status: Preliminary; translated from GB/EMBL/DDJB

A;Molecule type: DNA

A;Residues: 1-279 <WIL>

A;Cross-references: EMBL:Z67995; PIDN:CAA91943.1; GSPDB:GN00028; CESP:M153.1

A;Experimental source: clone M153

C;Genetics:

A;Gene: CESP:M153.1

A;Map position: X

A;Introns: 23/1; 106/3; 177/3; 248/3

C;Superfamily: Pyrroline-5-carboxylate reductase

Query Match 32.7%; Score 512.5; DB 1; Length 276;

Best Local Similarity 42.5%; Pred. No. 2.1e-33;

Matches 113; Conservative 52; Mismatches 90; Indels 11; Gaps 4;

QY 3 VGFITGAGOLAYRFTA---TAGIUSAHKTIASSEMNLPTVSALRKMGVNLTNSKETV 56

Db 13 VGFITGAGKMASTARGVVAASGVLPNRLCTAVHSNLRNDRVFESEFGVNPVFTSEEVKE 71

QY 59 SDVFLFLAVKPVHVKVASEIAPALKERHLYWNLSTALGITIRNIESLL--PTKSRVVR 117

Db 72 SDVFLFLAVKPVHVKVASEIAPALKERHLYWNLSTALGITIRNIESLL--PTKSRVVR 127

QY 119 NTPVWVQEGATVYVGTGHALVEGQLELQMLSSVGCFTEVEEDLIDAVTGLGSGPAPV 176

Db 128 NTPAAVGEAASVMSLGTGATEEDGAIWALFGAVGKTLKADKMFDAVTGLGSGSPAYF 187

QY 179 MALD-ADGGVKMGLPRLAQLQGAQALGAQMLDSEHQPCQQLDKWCSPGGATH 236

Db 188 LATEALAGGVAAGLPLRELAISASQTVGAATMSVTKHGPVLKDVTSPGGTTAGV 247

QY 237 HFLESGGFRLSLINAVASCIRTEL 262

Db 248 HELEKGSFRATLMAVVAAKRSREL 273

RESULT 4

S10186 Pyrroline-5-carboxylate reductase (EC 1.5.1.2) - soybean

C;Species: Glycine max (soybean)

C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999

C;Accession: S10186

R;Delauney, A. J.; Verma, D. P. S.

Mol. Gen. Genet. 221, 299-305, 1990

A;Title: A soybean gene encoding Delta-5-pyrroline-5-carboxylate reductase was isolated

A;Reference number: S10186; MUID:90340278

A;Accession: S10186

A;Molecule type: mRNA

A;Residues: 1-274

A;Cross-references: EMBL:XI6352; NID:918723; PIDN:CAA34401.1; PID:918724

A;Note: the authors translated the codon GAA for residue 115 as Pro and GTG for residue 116

C;Superfamily: Pyrroline-5-carboxylate reductase

C;Keywords: oxidoreductase; Proline biosynthesis

Query Match 32.0%; Score 501.5; DB 1; Length 274;

Best Local Similarity 41.5%; Pred. No. 1.5e-32;

Matches 111; Conservative 51; Mismatches 94; Indels 11; Gaps 4;

QY 2 SVGETGAGOLAYRFTA---TAGIUSAHKTIASSEMNLPTVSALRKMGVNLTNSKETV 57

Db 12 TLGFIGAGKMASTARGVVAASGVLPNRLCTAVHSNLRNDRVFESEFGVNPVFTSEEVKE 70

QY 58 HDVFLFLAVKPVHVKVASEIAPALKERHLYWNLSTALGITIRNIESLL--PTKSRVVR 117

Db 71 ESDVFLFLAVKPVHVKVASEIAPALKERHLYWNLSTALGITIRNIESLL--PTKSRVVR 126

QY 119 TNPVWVQEGATVYVGTGHALVEGQLELQMLSSVGCFTEVEEDLIDAVTGLGSGPAPV 176

Db 127 PNTPAAVGEAASVMSLGTGATEEDGAIWALFGAVGKTLKADKMFDAVTGLGSGSPAYF 186

QY 178 FMFLAD-ADGGVKMGLPRLAQLQGAQALGAQMLDSEHQPCQQLDKWCSPGGATHA 235

Db 187 YLATEALAGGVAAGLPLRELAISASQTVGAATMSVTKHGPVLKDVTSPGGTTAGV 246

QY 236 LHFLESGGFRLSLINAVASCIRTEL 262

Db	247	IHELENGFGFTLMLNAVVAAKRSREL	273	Matches 108; Conservative 38; Mismatches 89; Indels 34; Gaps 5;
RESULT	5			
T06477				
probable pyrroline-5-carboxylate reductase (EC 1.5.1.2) - garden pea				
C;Species: Pisum sativum (garden pea)				
C;Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 11-Jun-1999				
C;Accession: T06477				
R;Williamson, C.L.; Slocum, R.D.				
PInt. Physiol. 100: 1464-1470, 1992				
A;Title: Molecular cloning and evidence for osmoregulation of the di-pyrroline-5-carboxylic acid reductase number: T06477				
A;Status: preliminary; translated from GB/EMBL/DDJB				
A;Molecule type: mRNA				
A;Residues: 1-273 <WIL>				
A;Cross-references: EMBL:X62842; NID:920850; PIDN:CAA44646.1; PID:920851				
A;Experimental source: cv. Wando				
A;Gene: proc				
C;Genetics:				
C;Superfamily: pyrroline-5-carboxylate reductase				
C;Keywords: oxidoreductase; proline biosynthesis				
Query Match	30.6%	Score 479; DB 2; Length 273;		
Best Local Similarity	40.1%	Pred. No. 9.3e-31; Mismatches 5; Indels 5; Gaps 5;		
Matches 108; Conservative 55; Mismatches 90; Indels 16; Gaps				
Qy 2 SVGFTGAGQQLAYRF---TAGIISAHKITA--SSPEMNTPTVSALKMCGVNLTRSNKET 55				
Db 12 TLGFAGKAGMAESTIAGKASSGVLPSRRIVTAHSNPSRR---AAFESIGITVLSSNDV 67				
Qy 56 VKHSDVLFVAKPHPTIDEAGDVOARHIVVSCAGYTTSSVEKKLAQFQPAVKR 115				
Db 68 VRASNWVWVSVKPOLVKDVLKPLTOKRKLIVSVAAGKL---KDLQEWAGHERFIR 123				
Qy 116 CMTNTPVWVQEGATWYATGTHALVEDQOLLEQMLSSVGCETEVEEDLIDAVTGLSGSPA 175				
Db 124 VMPNTPAVQQAASVMSLGCATEEDBANLISOLFGSGIKIWKADDKFDDAATGGLSGSPA 183				
Qy 176 YAFMADL- ADGGVKGMLRRLAQIQLGAQMLLGAQMLLSEQHPCQLKDNVCSPGGATT 233				
Db 184 YIYLALTEALADGGVAGLBDRALSLASQVTLGAASMATLSKGHPQQLKDVTSPGGTT 243				
Qy 234 HALPLESGGGFRSLINAVBASCIRTEL 262				
Db 244 AGVHBLEKGFFRGTLMLNAVVAAKRSREL 272				
RESULT	6			
T36286				
pyrroline-5-carboxylate reductase - Streptomyces coelicolor				
C;Species: Streptomyces coelicolor				
C;Accession: T36286				
C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 21-Jan-2000				
R;Seeger, K.J.; Harris, D.; Bentley, S.D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.				
A;Subtilized to the EMBL Data Library, May 1998				
A;Reference number: 221603				
A;Accession: T36286				
A;Status: preliminary; translated from GB/EMBL/DDJB				
A;Molecule type: DNA				
A;Residues: 1-284 <SEER>				
A;Cross-references: EMBL:AL049819; PIDN:CAB42663.1; GSPDB:GN00070; SCOECDB:SCE7.04C				
A;Experimental source: strain A3(2)				
C;Genetics:				
C;Gene: proc				
C;Superfamily: pyrroline-5-carboxylate reductase				
C;Keywords: oxidoreductase; proline biosynthesis				
Query Match	26.9%	Score 421; DB 2; Length 267;		
Best Local Similarity	38.5%	Pred. No. 3.7e-26; Mismatches 96; Indels 26; Gaps 8;		
Matches 106; Conservative 47; Mismatches 96; Indels 26; Gaps 8;				
Qy 1 MSVGFPTGAGOLA---YRFTAGILSAHKITA--SSPEMNLPTVSALKMCGVNLTR 51				
Db 3 IQLGITIGGGVMAEALIARLIAEKTWAAPEEITVGEPEH---GARRYLOKTYQVRSPD 56				
Qy 52 NKETVHSDVLFVAKPHPTIDEAGDVOARHIVVSCAGYTTSSVEKKLAQFQPAK 111				
Db 57 NOEANAVSEVILLAVLKVPOVLDVLVLASSLAGGAN-RPLVISITLAGVSLQRIQKF---PDH 111				
Qy 112 KVIRCHNTPVWVQEGATWYATGTHALVEDGOL- L-EQLMSSVGFTEVEEDLIDAVTGL 169				
Db 112 ATTRAMHNTPTVW---GAGMATAAANKVNPDLAKAIFSAVGNVNPENMDAVTG 169				
Qy 170 SGSGPAYAFMADL- ADGGVKGMLRRLAQIQLGAQMLLSEQHPCQLKDNVC 227				
Db 170 SGSGPAYVALMTEALADGGVLAGLPPRAIAQKLAQVGLTAKLIKETEEHAPQIKVTS 229				
Qy 228 PGGATIHALHFLESSGGFRSLINAVBASCIRTEL 262				
Db 230 PGGTIAGVAVLEKMGFRSASIAEVAYRARSQEL 264				

RESULT 8

H85533 Pyrroline-5-carboxylate reductase [imported] - Escherichia coli (strain O157:H7)

C;Species: Escherichia coli

C;Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 31-Mar-2001

C;Accession: H85533

R;Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew, L.; Grobebeck, E.J.; Davis, N.W.; Lin, A.; Dimambata, E.; Potamitis, K.; Apodaca, N.; Nature 409, 529-533, 2001

A;Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.

A;Reference number: A85480; MUID:21074935; PMID:1120551

A;Accession: H85533

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-269 <STO>

A;Cross-references: GB:AB005174; NID:912513226; PIDN:AAG54732.1; GSPDB:GN00145; UWGP:Z04

C;Genetics: proc

C;Function: catalyzes reduction of Pyrroline-5-carboxylate to proline

A;Description: catalyzes reduction of Pyrroline-5-carboxylate to proline

A;Pathway: proline biosynthesis

R;Note: third enzyme in the proline biosynthetic Pathway

C;Superfamily: pyrroline-5-carboxylate reductase

C;Keywords: oxidoreductase; proline biosynthesis

C;Superfamily: pyrroline-5-carboxylate reductase

Query Match 25.9%; Score 406.5; DB 2; Length 269; Best Local Similarity 35.9%; Pred. No. 5.3e-25; Matches 99; Conservative 51; Mismatches 107; Indels 19; Gaps 6;

Qy 3 VGFIGAGOLAYVFTAAIGILSAHKITASSEPMNLPT--VSALR-KMGVNLTNSNKEVKKH 58

Db 5 IFIGCGGNMG-KAIGGLGLASQVLPQGQIWWYTPSPDKVAAHLIDKFGINAESAQEVQI 63

Qy 59 SDVLFVAKPHITPILDEGADQVARHIVWSCAGWVSSVEKKLMAFQPAKVKRCMT 118

Db 64 ADIFAAVKGIMKVLSEITSSLNKDSLWVSIAGVTLQDAR--ALGHDRKTIKAMP 120

Qy 119 NTPVWVVOEGATVYATGTHALVEDGOLLEQOLMSVGFCTEVEEDLIDAVGLSGSGPAYF 178

Db 121 NTPALVNAGMTSNTVNALVTPEDADVLNIFRCFGEAEVIAEPNIPVWVGSQSSPPAVF 180

Qy 179 MALD--ADGGVKGMLPPLRILATOLGAQALGAAKMLDSBQHPOPKDNYCSPGATIHL 236

Db 181 MFTTEAMADAAVLGGMPPRAQYKAQAVMGSAKAVLGETGHPGALKDMVCSPGTTIARV 240

Qy 237 HELESGGRSLLINAVEASCIRTELOSMADOKIS 272

Db 241 RVLEEKGFRAVIEAM-----TKCMERSEKLS 267

RESULT 9

RDCC

Pyrroline-5-carboxylate reductase (EC 1.5.1.2) - Escherichia coli

C;Species: Escherichia coli

C;Date: 13-Jun-1983 #sequence_revision 13-Jun-1983 #text_change 11-Jun-1999

C;Accession: J00385; B6467

R;Deutch, A.H.; Smith, C.J.; Rushlow, K.E.; Kretschmer, P.J.

Nucleic Acids Res. 10, 7701-7714, 1982

A;Title: Escherichia coli delta(1)-Pyrroline-5-carboxylate reductase: gene sequence, proc

A;Reference number: A00385; MUID:83116986

A;Accession: A00385

A;Molecule type: DNA

A;Residues: 1-269 <DBU>

A;Cross-references: GB:J01665; NID:9147358; PIDN:AAA8433.1; PID:9147359

A;Note: parts of this sequence, including the amino and carboxyl ends of the mature protein

A;Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; Co

A;Rose, D.J.; Mau, B.; Shao, Y.

Science 277, 1453-1462, 1997

A;Title: The complete genome sequence of Escherichia coli K-12.

A;Reference number: A64720; MUID:97426617

A;Accession: B64767

A;Status: nucleic acid sequence not shown; translation not shown

A;Molecule type: DNA

A;Residues: 1-269 <BLT>

A;Cross-references: GB:AE000145; GB:U00096; NID:91786580; PIDN:AACT73489.1; PID:91786585;

A;Experimental source: strain K-12, substrain MG1655

Query Match 24.1%; Score 377; DB 2; Length 273; Best Local Similarity 36.6%; Pred. No. 1.2e-22;

C;Genetics: proc

A;Gene: proc

A;Map position: 9 min

C;Function:

A;Description: catalyzes reduction of Pyrroline-5-carboxylate to proline

A;Pathway: proline biosynthesis

R;Note: third enzyme in the proline biosynthetic Pathway

C;Superfamily: pyrroline-5-carboxylate reductase

C;Keywords: oxidoreductase; proline biosynthesis

Query Match 25.7%; Score 402.5; DB 1; Length 269; Best Local Similarity 35.5%; Pred. No. 1.1e-24; Matches 98; Conservative 52; Mismatches 107; Indels 19; Gaps 6;

Qy 3 VGFIGAGOLAYVFTAAIGILSAHKITASSEPMNLPT--VSALR-KMGVNLTNSNKEVKKH 58

Db 5 IFIGCGGNMG-KAIGGLGLASQVLPQGQIWWYTPSPDKVAAHLIDKFGINAESAQEVQI 63

Qy 59 SDVLFVAKPHITPILDEGADQVARHIVWSCAGWVSSVEKKLMAFQPAKVKRCMT 118

Db 64 ADIFAAVKGIMKVLSEITSSLNKDSLWVSIAGVTLQDAR--ALGHDRKTIKAMP 120

Qy 119 NTPVWVVOEGATVYATGTHALVEDGOLLEQOLMSVGFCTEVEEDLIDAVGLSGSGPAYF 178

Db 121 NTPALVNAGMTSNTVNALVTPEDADVLNIFRCFGEAEVIAEPNIPVWVGSQSSPPAVF 180

Qy 179 MALD--ADGGVKGMLPPLRILATOLGAQALGAAKMLDSBQHPOPKDNYCSPGATIHL 236

Db 181 MFTTEAMADAAVLGGMPPRAQYKAQAVMGSAKAVLGETGHPGALKDMVCSPGTTIARV 240

Qy 237 HELESGGRSLLINAVEASCIRTELOSMADOKIS 272

Db 241 RVLEEKGFRAVIEAM-----TKCMERSEKLS 267

RESULT 10

J00418 Pyrroline-5-carboxylate reductase (EC 1.5.1.2) precursor - Pseudomonas aeruginosa (strain PA01)

C;Species: Pseudomonas aeruginosa

C;Date: 31-Dec-1991 #sequence_revision 31-Dec-1991 #text_change 31-Dec-2000

C;Accession: J00418; H83597

R;Savioz, A.; Jeenes, D.J.; Kocher, H.P.; Haas, D.

Gene 66, 107-111, 1990

A;Title: Comparison of proc and other housekeeping genes of Pseudomonas aeruginosa with other Pseudomonas aeruginosa strains

A;Reference number: J00418; MUID:90185238

A;Accession: J00418

A;Molecule type: DNA

A;Residues: 1-273 <SAV>

A;Cross-references: GB:M3557; NID:9151509; PIDN:AA25975.1; PID:9151510

A;Experimental source: strain PA01

A;Note: part of this sequence, including the amino end of the mature protein, was constructed by sequencing the cDNA clone pPA01-1.

R;Stover, C.K.; Pham, X.O.; Erwin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.; Adamian, S.; Yuan, Y.; Brody, D.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lory, S.; Olson, M.V.

Nature 406, 959-964, 2000

A;Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic pathogen

A;Reference number: A82950; MUID:20437337

A;Accession: H83597

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-273 <STO>

A;Cross-references: GB:AE004476; GB:AE004091; NID:99946234; PIDN:AAG03782.1; GSPDB:GN

A;Experimental source: strain PA01

A;Genetics: proc

A;Gene: PA0393

C;Superfamily: pyrroline-5-carboxylate reductase

C;Keywords: oxidoreductase; proline biosynthesis

F;2-273/Product: Pyrroline-5-carboxylate reductase #status experimental <MAT>

R. Kunst, F.; Ogasawa, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Berteletti, C.; Bron, S.; Brouillet, S.; Bruschi, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Chodat, P.; Ehrlich, S.D.; Emerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari, B.; Nature 390, 249-256, 1997
 A; Authors: Boulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Galler, J.; Harwood, C.R.; Renaut, A.; Hilbert, H.; Hosapelli, S.; Hosono, S.; Hulio, M.; Ichach, J.; Koettler, P.; Koningstein, G.; Krogh, S.; Kumano, M.; Kurappa, K.; Lapidus, A.; Lardinois, A.; Koettler, P.; Koningstein, G.; Krogh, S.; Kumano, M.; Kurappa, K.; Lapidus, A.; Lardinois, A.; Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Mauel, Y.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetelle, D.; Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadaie, Y.; Sato, T.; Scanlon, A.; Authors: Schleich, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Seror, Akeuchi, M.; Tamakoshi, A.; Tanaka, T.; Terpstra, P.; Tognoni, A.; Tosato, V.; Uchiyama, T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida, K.; Authors: Yoshikawa, H.F.; Zunstein, E.; Yoshikawa, H.; Danchin, A.
 A; Title: The complete genome sequence of the Gram-positive bacterium *Bacillus subtilis*. A; accession: 66964
 A; molecule type: DNA
 A; residues: 1-278 <RNU>
 A; Cross-references: GB:29916; GB:AL009126; NID:92634723; PIDN:CAB14312.1; PID:92634815
 A; Experimental source: strain 168
 C; Genetics:
 A; Gene: yqj0
 A; Superfamily: pyrrole-5-carboxylate reductase
 C; Keywords: oxidoreductase

Query Match 20.7%; Score 324.5; DB 1; Length 278;
 Best Local Similarity 29.9%; Pred. 1.8e-18; Mismatches 88; Conservative 54; Indels 113; Gaps 39; Gaps 7; Matches 88; Mismatches 113; Indels 39; Gaps 7;

Qy 3 VGFIGAGQAYR---FIAAGLISAHKTIASSPEMLPTVSALKRGMVNLTNSNKETVKH 58
 Db 4 IGFYVGAGSMKAEMINGLILQSGTTRKPEHIIYITRSNDRERLIEKETYSVRPCRDNEFFH 63

Qy 59 SDVIFLAVKEP-----HIIPFLDTEGADWQARHIVWSAGVTTSSE---KKLMA 106
 Db 64 TDITLAFKPKDAESIDSIRPKD-----LVIISVLAGLTTEIOHYFGRLK-- 112

Qy 107 FOPAPKVIRCTNTPVVQEGATVYATGTHALVEDGOLQIMSSVGFCTEWEEELIDAV 166
 Db 113 -----AVIRMPNTSAIRKSAFGNSTEAKSKNDIAKALLETTIGDATEVVEERHLDVA 167

Qy 167 TGLSGSGPAPAFMALDA--DGGYKMGIDPRLATOLGAQALIGAKMILDSEQHPCOLKD 224
 Db 168 TAIAGSGSPAYVYRTEAMERAKAQVKGLDKEATAKALIQLTMAQATDMLQSQSKQPERLKE 227

Qy 225 VCPSPGATIHALFELFESGGFRRSLINAVEASCIRTRLEPKPAARK 278
 Db 228 ITSPGGTTEAGLRLAQLDSRFEAIIHCIEETAKRSAE---KEQFAGAMER 276

RESULT 15
 G83837 15
 pyrrole-5-carboxylate reductase BH1503 [imported] - *Bacillus halodurans* (strain C-1
 C; Species: *Bacillus halodurans*
 C; Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 31-Dec-2000
 C; Accession: G83837
 R; Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fuji, F.; H
 Nucleic Acids Res. 28, 4317-4331, 2000
 A; Title: Complete genome sequence of the alkaliphilic bacterium *Bacillus halodurans* a
 A; Reference number: A83650; MUID:20263314
 A; Status: preliminary
 A; Molecule type: DNA
 A; Residues: 1-279 <STO>
 A; Cross-references: GB:AP001512; GB:BA000004; NID:910174030; PIDN:BA005222.1; GSPDB:G
 A; Experimental source: strain C-125
 C; Genetics:
 A; Gene: BHI503
 C; Superfamily: pyrrole-5-carboxylate reductase

Query Match 20.6%; Score 323.5; DB 2; Length 279;
 Best Local Similarity 33.2%; Pred. 2.2e-18; Mismatches 91; Conservative 47; Indels 11; Gaps 5; Matches 91; Mismatches 125; Indels 11; Gaps 5;

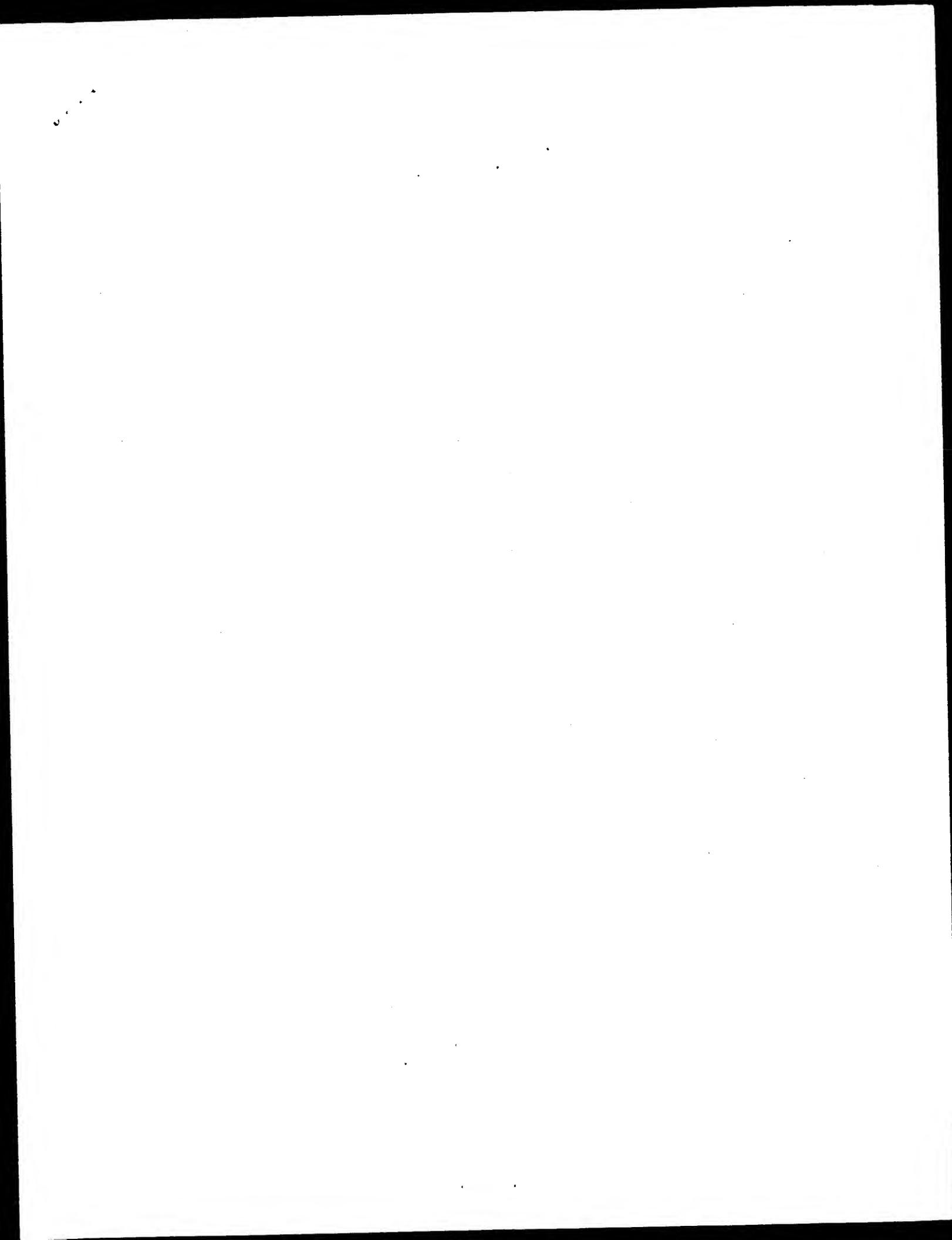
Qy 3 VGFIGAGQAYRTEAAGLISAHKTIASSPEMLPTVSALKRGMVNLTNSNKETVKH 57
 Db 7 ITFLGAGSM-ESTIAGLISKILLISQVIATNLCEAKLTKEEDRYGICITTOHQEA 65

Qy 58 HSDVIFLAVKEP-----HIIPFLDTEGADWQARHIVWSAGVTTSSE---KKLMA 117
 Db 66 QGTTIPLAKPKNTEATEIRGETEKQFLISVLAGTTSYIETLAHSPV--WVTRM 122

Db 118 TNTPVVQOCATVYATGTHALVEDGOLQIMSSVGFCTEWEEELIDAVGILSSGGPAA 177
 C; Species: *Thermus aquaticus*
 C; Date: 30-Sep-1993 #sequence_revision 20-Aug-1994 #text_change 26-Aug-1999
 C; Accession: JCC208
 R; Hoshino, T.; Kosuge, T.; Hidaka, Y.; Tabata, K.; Nakahara, T.
 Biochem. Biophys. Res. Commun. 199, 410-417, 1994
 A; Title: Molecular cloning and sequence analysis of the proc gene encoding delta-1-pyruvate kinase from *Thermus aquaticus*.
 A; Reference number: JCC208; MUID:94168609
 A; Molecule type: DNA
 A; Residues: 1-261 <HOS>
 C; Superfamily: pyrrole-5-carboxylate reductase
 C; Keywords: oxidoreductase

Query Match 20.6%; Score 323.5; DB 2; Length 261;
 Best Local Similarity 33.9%; Pred. No. 2e-18; Mismatches 45; Indels 29; Gaps 11; Matches 93; Conservative 45; Mismatches 107; Indels 29; Gaps 11;

Search completed: January 31, 2002, 14:54:14
 Job time: 44 sec



GenCore version 4.5
 Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model.

Run on:

January 31, 2002, 14:53:55 ; Search time 24.93 Seconds

(without alignments)

1842.339 Million cell updates/sec

Title: US-09-912-717-1

Perfect score: 1567

Sequence: 1 MSVFIGAGOLAYRFTAAGI..... TPSSPGKLRLSLALGGKKD 314

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched:

473505 seqs, 146272329 residues

Total number of hits satisfying chosen parameters: 473505

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- 1: SP TREMBL_17;*
- 2: SP archaea;*
- 3: SP bacteria;*
- 4: SP fungi;*
- 5: SP invertebrate;*
- 6: SP mammal;*
- 7: SP mhc;*
- 8: SP organelle;*
- 9: SP phage;*
- 10: SP plant;*
- 11: SP rodent;*
- 12: SP virus;*
- 13: SP vertebrate;*
- 14: SP unclassified;*

Pred. No. 1s is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1231.5	78.6	319	4 Q9Y5J4	09Y5J4 homo sapien
2	898	57.3	272	4 QHBBQ4	09Y5J4 homo sapien
3	578	56.9	279	5 Q21544	09Y5J4 homo sapien
4	531.5	34.3	274	4 Q9H896	Q21544 caenorhabditis elegans
5	522.5	33.3	274	11 Q9DC44	Q9H896 homo sapien
6	500.5	32.1	273	5 Q9VEJ3	Q9DC44 mus musculus
7	500.5	31.9	274	10 Q9AVM3	Q9VEJ3 drosophila melanogaster
8	481	30.7	280	5 Q9V3F8	Q9AVM3 vigna unguiculata
9	469	29.9	280	5 Q96643	Q9V3F8 drosophila melanogaster
10	456	29.1	270	1 Q9H999	Q96643 drosophila melanogaster
11	448	28.6	266	2 Q87725	Q9H999 methanosaeca
12	423	27.0	284	2 Q9X8G1	Q87725 clostridium
13	399	25.5	272	2 Q9K308	Q8X8G1 streptomyces
14	380	24.3	248	11 Q9D012	Q9K308 bacillus cereus
15	376	24.0	254	5 Q9NGS4	Q9D012 mus musculus
16	344	22.0	279	1 Q9HLT6	Q9NGS4 leishmania
17	325.5	20.8	273	2 Q9RAG8	Q9HLT6 thermoplasm
18	323.5	20.6	279	2 Q9KCR6	Q9RAG8 nostoc sp.
19	322	20.5	273	5 Q27722	Q9KCR6 bacillus haemolyticus

ALIGNMENTS

RESULT	1
ID	Q9Y5J4
AC	Q9Y5J4;
DT	01-NOV-1999 (TREMBL_1.12, Created)
DT	01-NOV-1999 (TREMBL_1.12, Last sequence update)
DT	01-JUN-2001 (TREMBL_1.12, Last annotation update)
DE	PYRROLINE 5-CARBOXYLATE REDUCTASE ISOFORM.
GN	P5CR2.
OS	Homo sapiens (Human).
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
OX	NCBI_TaxID:9606;
RN	[1]
RP	SEQUENCE FROM N.A.
RA	Kok L.D.S., Au N.T.C.C., Fung K.P., Tsui S.K.W., Lee C.Y., Wayne M.M.Y.; "Molecular characterization of a novel human pyrroline-5-carboxylate reductase isolated in liver." Submitted (MAY-1999) to the EMBL/genBank/DBJ databases.
RT	"Molecular characterization of a novel human pyrroline-5-carboxylate reductase isolated in liver." Submitted (MAY-1999) to the EMBL/genBank/DBJ databases.
RL	RT
DR	EMBL; AF151351; AD34611_1; -.
DR	InterPro; IPR000304; P5CR.
DR	Protein; PF01089; P5CR; 1.
DR	PROSITE; PS00521; P5CR; 1.
SEQ	SEQUENCE: 319 AA; 33726 MW; 01D49702A59883B CRC64;
Query	MSVFIGAGOLAYRFTAAGI..... TPSSPGKLRLSLALGGKKD
Match	MSVFIGAGOLAYRFTAAGI..... TPSSPGKLRLSLALGGKKD
Best Local Similarity	78.6%
Length	319
Mismatches	81.7%
Conservative	7
Mismatches	41
Indels	11
Gaps	4
Query Match	MSVFIGAGOLAYRFTAAGI..... TPSSPGKLRLSLALGGKKD
Best Local Similarity	78.6%
Length	319
Mismatches	81.7%
Conservative	7
Mismatches	41
Indels	11
Gaps	4
Query	MSVFIGAGOLAYRFTAAGI..... TPSSPGKLRLSLALGGKKD
Best Local Similarity	78.6%
Length	319
Mismatches	81.7%
Conservative	7
Mismatches	41
Indels	11
Gaps	4
Query	MSVFIGAGOLAYRFTAAGI..... TPSSPGKLRLSLALGGKKD
Best Local Similarity	78.6%
Length	319
Mismatches	81.7%
Conservative	7
Mismatches	41
Indels	11
Gaps	4
Query	MSVFIGAGOLAYRFTAAGI..... TPSSPGKLRLSLALGGKKD
Best Local Similarity	78.6%
Length	319
Mismatches	81.7%
Conservative	7
Mismatches	41
Indels	11
Gaps	4
Query	MSVFIGAGOLAYRFTAAGI..... TPSSPGKLRLSLALGGKKD
Best Local Similarity	78.6%
Length	319
Mismatches	81.7%
Conservative	7
Mismatches	41
Indels	11
Gaps	4
Query	MSVFIGAGOLAYRFTAAGI..... TPSSPGKLRLSLALGGKKD
Best Local Similarity	78.6%
Length	319
Mismatches	81.7%
Conservative	7
Mismatches	41
Indels	11
Gaps	4
Query	MSVFIGAGOLAYRFTAAGI..... TPSSPGKLRLSLALGGKKD
Best Local Similarity	78.6%
Length	319
Mismatches	81.7%
Conservative	7
Mismatches	41
Indels	11
Gaps	4
Query	MSVFIGAGOLAYRFTAAGI..... TPSSPGKLRLSLALGGKKD
Best Local Similarity	78.6%
Length	319
Mismatches	81.7%
Conservative	7
Mismatches	41
Indels	11
Gaps	4
Query	MSVFIGAGOLAYRFTAAGI..... TPSSPGKLRLSLALGGKKD
Best Local Similarity	78.6%
Length	319
Mismatches	81.7%
Conservative	7
Mismatches	41
Indels	11
Gaps	4
Query	MSVFIGAGOLAYRFTAAGI..... TPSSPGKLRLSLALGGKKD
Best Local Similarity	78.6%
Length	319
Mismatches	81.7%
Conservative	7
Mismatches	41
Indels	11
Gaps	4
Query	MSVFIGAGOLAYRFTAAGI..... TPSSPGKLRLSLALGGKKD
Best Local Similarity	78.6%
Length	319
Mismatches	81.7%
Conservative	7
Mismatches	41
Indels	11
Gaps	4
Query	MSVFIGAGOLAYRFTAAGI..... TPSSPGKLRLSLALGGKKD
Best Local Similarity	78.6%
Length	319
Mismatches	81.7%
Conservative	7
Mismatches	41
Indels	11
Gaps	4
Query	MSVFIGAGOLAYRFTAAGI..... TPSSPGKLRLSLALGGKKD
Best Local Similarity	78.6%
Length	319
Mismatches	81.7%
Conservative	7
Mismatches	41
Indels	11
Gaps	4
Query	MSVFIGAGOLAYRFTAAGI..... TPSSPGKLRLSLALGGKKD
Best Local Similarity	78.6%
Length	319
Mismatches	81.7%
Conservative	7
Mismatches	41
Indels	11
Gaps	4
Query	MSVFIGAGOLAYRFTAAGI..... TPSSPGKLRLSLALGGKKD
Best Local Similarity	78.6%
Length	319
Mismatches	81.7%
Conservative	7
Mismatches	41
Indels	11
Gaps	4
Query	MSVFIGAGOLAYRFTAAGI..... TPSSPGKLRLSLALGGKKD
Best Local Similarity	78.6%
Length	319
Mismatches	81.7%
Conservative	7
Mismatches	41
Indels	11
Gaps	4
Query	MSVFIGAGOLAYRFTAAGI..... TPSSPGKLRLSLALGGKKD
Best Local Similarity	78.6%
Length	319
Mismatches	81.7%
Conservative	7
Mismatches	41
Indels	11
Gaps	4
Query	MSVFIGAGOLAYRFTAAGI..... TPSSPGKLRLSLALGGKKD
Best Local Similarity	78.6%
Length	319
Mismatches	81.7%
Conservative	7
Mismatches	41
Indels	11
Gaps	4
Query	MSVFIGAGOLAYRFTAAGI..... TPSSPGKLRLSLALGGKKD
Best Local Similarity	78.6%
Length	319
Mismatches	81.7%
Conservative	7
Mismatches	41
Indels	11
Gaps	4
Query	MSVFIGAGOLAYRFTAAGI..... TPSSPGKLRLSLALGGKKD
Best Local Similarity	78.6%
Length	319
Mismatches	81.7%
Conservative	7
Mismatches	41
Indels	11
Gaps	4
Query	MSVFIGAGOLAYRFTAAGI..... TPSSPGKLRLSLALGGKKD
Best Local Similarity	78.6%
Length	319
Mismatches	81.7%
Conservative	7
Mismatches	41
Indels	11
Gaps	4
Query	MSVFIGAGOLAYRFTAAGI..... TPSSPGKLRLSLALGGKKD
Best Local Similarity	78.6%
Length	319
Mismatches	81.7%
Conservative	7
Mismatches	41
Indels	11
Gaps	4
Query	MSVFIGAGOLAYRFTAAGI..... TPSSPGKLRLSLALGGKKD
Best Local Similarity	78.6%
Length	319
Mismatches	81.7%
Conservative	7
Mismatches	41
Indels	11
Gaps	4
Query	MSVFIGAGOLAYRFTAAGI..... TPSSPGKLRLSLALGGKKD
Best Local Similarity	78.6%
Length	319
Mismatches	81.7%
Conservative	7
Mismatches	41
Indels	11
Gaps	4
Query	MSVFIGAGOLAYRFTAAGI..... TPSSPGKLRLSLALGGKKD
Best Local Similarity	78.6%
Length	319
Mismatches	81.7%
Conservative	7
Mismatches	41
Indels	11
Gaps	4
Query	MSVFIGAGOLAYRFTAAGI..... TPSSPGKLRLSLALGGKKD
Best Local Similarity	78.6%
Length	319
Mismatches	81.7%
Conservative	7
Mismatches	41
Indels	11
Gaps	4
Query	MSVFIGAGOLAYRFTAAGI..... TPSSPGKLRLSLALGGKKD
Best Local Similarity	78.6%
Length	319
Mismatches	81.7%
Conservative	7
Mismatches	41
Indels	11
Gaps	4
Query	MSVFIGAGOLAYRFTAAGI..... TPSSPGKLRLSLALGGKKD
Best Local Similarity	78.6%
Length	319
Mismatches	81.7%
Conservative	7
Mismatches	41
Indels	11
Gaps	4
Query	MSVFIGAGOLAYRFTAAGI..... TPSSPGKLRLSLALGGKKD
Best Local Similarity	78.6%
Length	319
Mismatches	81.7%
Conservative	7
Mismatches	41
Indels	11
Gaps	4
Query	MSVFIGAGOLAYRFTAAGI..... TPSSPGKLRLSLALGGKKD
Best Local Similarity	78.6%
Length	319
Mismatches	81.7%
Conservative	7
Mismatches	41
Indels	11
Gaps	4
Query	MSVFIGAGOLAYRFTAAGI..... TPSSPGKLRLSLALGGKKD
Best Local Similarity	78.6%
Length	319
Mismatches	81.7%
Conservative	7
Mismatches	41
Indels	11
Gaps	4
Query	MSVFIGAGOLAYRFTAAGI..... TPSSPGKLRLSLALGGKKD
Best Local Similarity	78.6%
Length	319
Mismatches	81.7%
Conservative	7
Mismatches	41
Indels	11
Gaps	4
Query	MSVFIGAGOLAYRFTAAGI..... TPSSPGKLRLSLALGGKKD
Best Local Similarity	78.6%
Length	319
Mismatches	81.7%
Conservative	7
Mismatches	41
Indels	11
Gaps	4
Query	MSVFIGAGOLAYRFTAAGI..... TPSSPGKLRLSLALGGKKD
Best Local Similarity	78.6%
Length	319
Mismatches	81.7%
Conservative	7
Mismatches	41
Indels	11
Gaps	4
Query	MSVFIGAGOLAYRFTAAGI..... TPSSPGKLRLSLALGGKKD
Best Local Similarity	78.6%
Length	319
Mismatches	81.7%
Conservative	7
Mismatches	41
Indels	11
Gaps	4
Query	MSVFIGAGOLAYRFTAAGI..... TPSSPGKLRLSLALGGKKD
Best Local Similarity	78.6%
Length	319
Mismatches	81.7%
Conservative	7
Mismatches	41
Indels	11
Gaps	4
Query	MSVFIGAGOLAYRFTAAGI..... TPSSPGKLRLSLALGGKKD
Best Local Similarity	78.6%
Length	319
Mismatches	81.7%
Conservative	7
Mismatches	41
Indels	11
Gaps	4
Query	MSVFIGAGOLAYRFTAAGI..... TPSSPGKLRLSLALGGKKD
Best Local Similarity	78.6%
Length	319
Mismatches	81.7%
Conservative	7
Mismatches	41
Indels	11
Gaps	4
Query	MSVFIGAGOLAYRFTAAGI..... TPSSPGKLRLSLALGGKKD
Best Local Similarity	78.6%
Length	319
Mismatches	81.7%
Conservative	7
Mismatches	41
Indels	11
Gaps	4
Query	MSVFIGAGOLAYRFTAAGI..... TPSSPGKLRLSLALGGKKD
Best Local Similarity	78.6%
Length	319
Mismatches	81.7%
Conservative	7
Mismatches	41
Indels	11
Gaps	4
Query	MSVFIGAGOLAYRFTAAGI..... TPSSPGKLRLSLALGGKKD
Best Local Similarity	78.6%
Length	319
Mismatches	81.7%
Conservative	7
Mismatches	41
Indels	11
Gaps	4
Query	MSVFIGAGOLAYRFTAAGI..... TPSSPGKLRLSLALGGKKD
Best Local Similarity	78.6%
Length	319
Mismatches	81.7%
Conservative	7
Mismatches	41
Indels	11
Gaps	4
Query	MSVFIGAGOLAYRFTAAGI..... TPSSPGKLRLSLALGGKKD
Best Local Similarity	78.6%
Length	319
Mismatches	81.7%
Conservative	7
Mismatches	41
Indels	11
Gaps	4
Query	MSVFIGAGOLAYRFTAAGI..... TPSSPGKLRLSLALGGKKD
Best Local Similarity	78.6%
Length	319
Mismatches	81.7%
Conservative	7
Mismatches	41
Indels	11
Gaps	4
Query	MSVFIGAGOLAYRFTAAGI..... TPSSPGKLRLSLALGGKKD
Best Local Similarity	78.6%
Length	319
Mismatches</	

RT "NEDO human cDNA sequencing project.";
 RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AK023914; BAB14721.1; -;
 DR InterPro: IPR00304; P5CR.
 DR Pfam: PF01089; P5CR; 1.
 DR PROTEIN: PS00521; P5CR; 1.
 SQ SEQUENCE 274 AA; 28663 MW; 846FDEC603F3B540 CRC64;
 .
 Query Match 34.3%; Score 537.5; DB 4; Length 274;
 Best Local Similarity 45.3%; Pred. No. 5.1e-35; Indels 13; Gaps 5;
 Matches 121; Conservative 45%; Mismatches 88; Indels 13; Gaps 5;
 .
 QY 3 VGRIGAGO---LAYRTAAIGLISANKIIASSP-EMNLQPTVSALRKMGVNLTNSRNKETVK 57
 ||||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:
 Db 11 VGFVGAGRMAGATAQGLIRAGKVEAQHILASPTDNRUCHQAL--GCRTHSNOEVQ 67
 .
 QY 58 HSIYLFVFLAVKPHITPFLIDETGQDQARHIVVSCAGVTISVEKKLMAFQAPKVIRCM 117
 |||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:
 Db 68 SCLIVVIFATKPHVLPVLAEVAVPVTEHILVSVAGVSLSTLELL--PPNTRVLRLV 124
 .
 QY 118 TNTPVWQOEGATVATGTHALVEDGQLEQMLQSSVGCTEFEEDLDAVGLGSPPAY- 176
 |||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:
 Db 125 PNLPCCVWQEGATVIMARHGHHVGSSSETKUQHILAECCGEEPEAYVDTGJGSGVAFV 184
 .
 QY 177 -AFMALDADGGYKMGPLRRLATOLQGAQLLGAKAAMLDSEQHPCQLEKDNVCSPPGATHA 235
 |||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:
 Db 185 CARSEALAEAGAVKNGIMPSSLAIRIAQTLGTAQMLLHEGQHQAOLRSVDCTPGGTIYG 244
 .
 QY 236 LHFILESGGRFLSLINAVEASCIRTEL 262
 |||:|||:|||:|||:|||:|||:|||:
 Db 245 LHALEQGLRAATMSAVEATCRAKEL 271
 .
 RESULT 5
 Q9DC4 PRELIMINARY; PRT: 274 AA.
 AC Q9DC4;
 DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
 DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
 DE 11.0058B13RIK PROTEIN.
 GN MUS_musculus (Mouse)
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=57BL/6J; TISSUE=KIDNEY;
 RX MEDLINE=21085660; PubMed=1217851;
 RA Kawai J., Shinagawa A., Shibusawa K., Yoshino M., Itoh M., Ishii Y.,
 RA Arakawa T., Hara A., Fukunishi Y., Kono H., Adachi J., Fukuda S., Yamamoto T.,
 RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamamoto T.,
 RA Saito T., Matsuda H.A., Gojobori T., Bono H., Kasukawa T., Saito R.,
 RA Kado K., Matsuda H.A., Ashburner M., Battalov S., Casavant T.,
 RA Fleischmann W., Gaestelmann T., Gissi C., King B., Kochiwa H.,
 RA Kuehl P., Lewis S., Matsuo Y., Nikaido T., Peso G., Quackenbush J.,
 RA Schriml L.M., Staubli R., Suzuki Y., Tomita M., Wagner L., Washio T.,
 RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
 RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.,
 RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Garibaldi M.,
 RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya J., Lee N.H.,
 RA Lyons P., Marchlioni L., Maslina J., Mazzarelli J., Mombretti P.,
 RA Nordone P., Ring B., Ringwald M., Rodriguez T., Sakamoto N.,
 RA Sasaki H., Sato K., Schoenbach C., Seva T., Shibata Y., Storch K.-F.,
 RA Wynnshaw-Borisi A., Yoshida H., Hasegawa Y., Kawaji H., Kohetsuki S.,
 RA Hayashiaki Y.;
 RT Functional annotation of a full-length mouse cDNA collection.";
 RL Nature 405:685-690 (2000).
 DR EMBL: AK002912; BAB22451.1; -;
 DR MGD: MGI:191344; 1110058B13RIK.
 .
 RESULT 6
 Q9VEJ3 PRELIMINARY; PRT: 273 AA.
 AC Q9VEJ3;
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
 DT 01-MAY-2000 (TREMBLrel. 13, Last annotation update)
 DE CG5840 PROTEIN.
 GN CG5840.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydriidae; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BERKELEY;
 RX MEDLINE=20196006; PubMed=10731132;
 RA Adams M.D., Celminker S.E., Holt R.A., Evans C.A., "ayne J.D.,
 RA Ananatades P.G., Scherer S.E., Li P.W., Hoskin: "..., Galie R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner: "..., Henderson S.N.,
 RA Suttorp G.G., Wormtan J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Wortsman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Wan K.H., Doyle C., Baxter E.G., Holt G., Nelson C.R., Miklos G.L.G.,
 RA Adril J.F., Agbayani A., An H.-J., Andress Pflanckoch C., Baldwin D.,
 RA Ballieu R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Beres P.V., Berman B.P., Bhandari D., Bolshakov S.,
 RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
 RA Burtis K.D., Butler H., Cadile E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Daveyport L.B., Davies P.,
 RA de Pablo B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Douc T., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin R.K.J., Evangelista C.C., Ferraz C., Ferreria S., Fleischmann W.,
 RA Fosler C., Gabrielle A.E., Garg N.S., Gelbart W.M., Glaser K.,
 RA Glode A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
 RA Jalali M., Kalush F., Karrien G.H., Ke Z., Keunison J.A., Kelchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattie B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Mishina N.V., Mobarry C., Morris J., Moslareff A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,

RA	Zheng X.-H., Zhong F.-N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O., Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.; "The genome sequence of <i>Drosophila melanogaster</i> "; <i>Science</i> 287:2185-2195(2000).
RL	SEQUENCE FROM N.A.
RC	STRAINeCS;
RA	Misener S.R., Walker V.K.; "High density of unrelated genes showing overlapping and introntronic transcription units in <i>Drosophila</i> "; Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
RT	EMBL: AE003724; AAH56265.1; EMBL: AE17029; AAH4940.1; FlyBase: FBgn0015781; P5CR; InterPro: IPR000304; P5CR.
RL	DR: PFM; PF01089; P5CR; 1; PROSITE; PS00521; P5CR; 1. KW: Oxidoreductase.
DR	SEQUENCE 280 AA; 29621 MW; D3D4FD9611574684 CRC64;
Query Match	29.9%; Score 469; DB 5; Length 280;
Best Local Similarity	38.4%; Pred. No. 1. 5e-29;
Matches	106; Conservative 51; Mismatches 97; Indels 22; Gaps 6;
Qy	3 VFGFIGAGOLAYR---TAAGILSAHKITASSEPMNLPTVSALRKMGVNLTRSNKEVKH 58
Db	8 LGFIGGGNMVAIGSGLVRCGIVKASQVSGP--HIEINQRWRLGAVTCDDNCVLEH 65
Qy	59 SDVLFELAVKPHIIPFILDEIGADVOARH-----VWSAGAVVLSSEKVKLMAFOP 109
Db	66 SDIVFICVKPH---MLTPCAAQKLYKHPASKDASKLVLVSLAGTSIQLTLEAFSFMGs 121
Qy	110 AP-KVTRCMNTPVVVOEGATVYATGTHALVEDQGLQELQMSVSGCTEVEDLDIATG 168
Db	122 SELKVIRTMNTSVOYGECTVYTGARNVSHIDELKTHLMVLNALGLAQVPEMSMTDAVG 181
Qy	169 LSGSGPAAVMAHD--ADGGVKGQLPRLAQGLAGLAKMILQDSEQHPCQKDNVC 226
Db	182 VAGGGPAVYITTEALADGGVKGQVPMQALQFAQTLGAAKTVLITGKHPAVLDEVc 241
Qy	227 SPGGATIHALHELSGGFERSLILNAVEASCIRTEL 262
Db	242 SPGGATIVGHELEKGNLRSTLINAEEKSSORSAEL 277
RESULT	10
ID	09HH9
AC	09HH99
PRELIMINARY	
PRT	270 AA.
096643	PRELIMINARY; PRT; 280 AA.
ID	096643; 01-MAY-1999 (TREMBlrel. 10, Last sequence update)
AC	096643; 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
DT	01-MAY-1999 (TREMBlrel. 10, Last sequence update)
DT	01-JUN-2001 (TREMBlrel. 17, Last annotation update)
DE	PROC.
GN	OS
OC	Methanobacteria acetivorans
OC	Archaea: Euryarchaeota; Methanobacteria; Methanobarcinales; Methanobarcinaceae; NCBI_TAXID=2214;
OX	[1]
RN	SEQUENCE FROM N.A.
RP	RC STRAINeC2A;
RA	Zhang J.-K., White A.K., Kueftner H.C., Boccazzi P., Metcalf W.W.; "Molecular and genetic analysis of proline biosynthesis in the methanogenic archaeon <i>Methanobarcina acetivorans</i> C2A"; Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.
RT	EMBL: AE305580; AAC22033.1; InterPro: IPR000304; P5CR.
DR	InterPro: IPR000304; P5CR; 1; PROSITE: PS00521; P5CR; 1.
DR	SEQUENCE 270 AA; 27943 MW; 50BC656AFCL0B1CF CRC64;
Query Match	29.1%; Score 456; DB 1; Length 270;
Best Local Similarity	39.7%; Pred. No. 1. 5e-28;
Matches	106; Conservative 52; Mismatches 97; Indels 12; Gaps 5;
Qy	3 VFGFIGAGOLAYR---TAAGILSAHKITASSEPMNLPTVSALRKMGVNLTRSNKEVKH 57
Db	6 IGFIGAGKNSAAMGTTKAGINPPENIGS--DWEPFLKDQKLGIVSTDNAIVR 63
Qy	58 HSDVLFELAVKPHIIPFILDEIGADVOARHIVVSCAAGTISSEKVKLMAFOPVKRCM 117
Db	64 ESDILFLAVKPKQTLSSVLSNLNEITSEKLVIAVPLSYEDALL--EGTRVVRM 120
Qy	118 TNTPVVVOEGATVYATGTHALVEDQGLQELQMSVSGCTEVEDLDIATGGLSPAYA 177
Db	121 PNTRATVSEASGIAQPKRNATPDEKLALEFSAVTAQVPESELMDAVGLSSGPAT 180
Qy	178 FMAHD--ADGGVKGQLPRLAQGLAGLAKMILQDSEQHPCQKDNVCSPGGATHA 235
Db	181 FPVTEAMDAGAVIEMGRKSALTAQOVTGAKMALEGMHPGELKDWTSPGTTG 240

Thu Jan 31 15:07:59 2002

us-09-912-717-1.rspt

Page 8

Db 110 EGVISITSGNSAVTPDBEAUVMKLFGAIGKAYLVAESAIIHGCVGVAGSSPAYVFMFEALS 169
QY 184 DGCVKMGLPRRLAIQLGAAQALGAAKMLLSEQHQHQCQLKVNVCPPGATHAIFPLECG 243
Db 170 DGAVRGGIPRAQSSEMAAQVILGAAKMLQESGKTPGALKDNVCSPGGTIEAVRPLECG 229
QY 244 FRSILINNAVEASCTRBLQSM 265
Db 230 MRSVIEAMIECMERKSEFEKI 251

Search completed: January 31, 2002, 14:57:19
Job time: 204 sec

Gencore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 31, 2002, 14:53:30 ; Search time 23.41 Seconds

(without alignments)
993.550 Million cell updates/sec

Title: US-09-912-717-1

Perfect score: 1

Sequence: MSVFIGAGOLAYRFTAAGI. TPSSPCKLRLSALAGGKKD 314

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 522463 seqs, 7407290 residues

Total number of hits satisfying chosen parameters: 522463

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0
Maximum Match 100%

Listing first 45 summaries

Database :

A_Geneseq_1101: * /SIDS8/gcadata/geneseq/geneseq/geneseq/AA1980.DAT: *

2: /SIDS8/gcadata/geneseq/geneseq/geneseq/AA1981.DAT: *

3: /SIDS8/gcadata/geneseq/geneseq/geneseq/AA1982.DAT: *

4: /SIDS8/gcadata/geneseq/geneseq/geneseq/AA1983.DAT: *

5: /SIDS8/gcadata/geneseq/geneseq/geneseq/AA1984.DAT: *

6: /SIDS8/gcadata/geneseq/geneseq/geneseq/AA1985.DAT: *

7: /SIDS8/gcadata/geneseq/geneseq/geneseq/AA1986.DAT: *

8: /SIDS8/gcadata/geneseq/geneseq/geneseq/AA1987.DAT: *

9: /SIDS8/gcadata/geneseq/geneseq/geneseq/AA1988.DAT: *

10: /SIDS8/gcadata/geneseq/geneseq/geneseq/AA1989.DAT: *

11: /SIDS8/gcadata/geneseq/geneseq/geneseq/AA1990.DAT: *

12: /SIDS8/gcadata/geneseq/geneseq/geneseq/AA1991.DAT: *

13: /SIDS8/gcadata/geneseq/geneseq/geneseq/AA1992.DAT: *

14: /SIDS8/gcadata/geneseq/geneseq/geneseq/AA1993.DAT: *

15: /SIDS8/gcadata/geneseq/geneseq/geneseq/AA1994.DAT: *

16: /SIDS8/gcadata/geneseq/geneseq/geneseq/AA1995.DAT: *

17: /SIDS8/gcadata/geneseq/geneseq/geneseq/AA1996.DAT: *

18: /SIDS8/gcadata/geneseq/geneseq/geneseq/AA1997.DAT: *

19: /SIDS8/gcadata/geneseq/geneseq/geneseq/AA1998.DAT: *

20: /SIDS8/gcadata/geneseq/geneseq/geneseq/AA1999.DAT: *

21: /SIDS8/gcadata/geneseq/geneseq/geneseq/AA2001.DAT: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1567	100.0	314	21	AAB20584
2	1509	96.3	320	22	AAB74779
3	1306	83.3	319	22	AAB93027
4	1045	66.7	212	22	AAB56640
5	6042	66.5	255	21	AAB92517
6	537.5	34.3	274	21	AAB92517
7	537.5	34.3	274	22	AAB25591
8	534.5	34.1	274	22	AAG66956
9	512.5	32.7	276	21	AAG06978
10	487.5	31.1	256	21	AAG06979
11	402.5	25.7	269	22	AAG64106

Arabidopsis thalia
C glutamicum prote
Corynebacterium gl
Human secreted pro
S. epidermidis ope
Arabidopsis thalia
CEP-B Homo sapi
C glutamicum prote
Corynebacterium gl
Corynebacterium gl
Glycerol-3 Phosph
S. aggregatum PKS
Enzyme involved in
Biosynthetic enzym
S. putrefaciens EP
S. putrefaciens PK
Stenotrophomonas putref
Arabidopsis thalia
Pancratium anemonia of
Glycerol-3-phosphate
Corynebacterium gl
Arabidopsis thalia
Arabidopsis thalia
S. putrefaciens ei
Caulobacter cresce
Caulobacter cresce
C glutamicum prote
Corynebacterium gl
Arabidopsis thalia
C glutamicum prote
Streptococcus pneu
Streptococcus pneu
Arabidopsis thalia
Arabidopsis thalia

ALIGNMENTS

RESULT	1
ID	AAB20584
ID	AAB20584 standard; Protein: 314 AA.
AC	
XX	
XX	AAB20584;
XX	
DT	13-DEC-2000 (first entry)
XX	
DE	Human delta-1-pyrroline-5-carboxylate reductase homologue protein.
XX	
XX	Human; delta-1-pyrroline-5-carboxylate reductase homologue; P5CRH; nootropic; neuroprotective; osteopathic; cytosolic; antiParkinsonian; cardiot; osteopathic; ophthalmological; hepatotropic; gene therapy; neuronal disorder; connective tissue disorder; cell proliferation; anesthesia; Alzheimer's disease; amnesia; dementia; Parkinson's disease; cystic fibrosis; myocardial fibrosis; osteoporosis; Martfan syndrome; cirrhosis; leukaemia; lymphoma; sarcoma; bone marrow cancer.
XX	
OS	Synthetic.
XX	
PN	US6100075-A.
XX	
PD	08-AUG-2000.
XX	
PP	18-JUN-1998; 98US-0099676.
XX	
PR	18-JUN-1998; 98US-0099676.
XX	
PA	(INCY-) INCYTE PHARM INC.
XX	
PI	Hillman JL, Baughn MR, Corley NC;
XX	
DR	WPI: 2000-548292/50.
DR	N-PSDB; AAB8073.

XX
PT New human delta 1-pyrroline-5-carboxylate reductase homolog (P5CRH) and
PT polynucleotides encoding P5CRH, useful for diagnosing, treating or
PT preventing neuronal or connective tissue disorders, or disorders of
PT cell proliferation
XX
PS Claim 1; Fig 1A-E; 32pp; English.

XX
The present sequence represents human delta 1-pyrroline-5-carboxylate
CC reductase homologue (P5CRH). P5CRH has nootropic, neuroprotective,
CC osteopathic, cyrostatic, anti-parkinsonian, cardiant, osteopathic,
CC ophthalmological and hepatotropic activities, and can be used in gene
CC therapy. P5CRH and the polynucleotides encoding it are useful for
CC diagnosing, treating or preventing neuronal disorders, connective tissue
CC disorders, or disorders of cell proliferation. These may be applied to
any subject, e.g. dogs, cats, cows, horses, rabbits, monkeys, or
CC preferably humans. The polynucleotide sequences may be used to detect
CC and quantitate gene expression in biopsied tissues where expression of
CC the polypeptide may correlate with the disease. The diagnostic assay may
CC be used to determine absence, presence, and excess expression of the
CC polypeptide, and to monitor regulation of the polypeptide levels during
CC therapeutic intervention. The polypeptide or its fragments are useful
CC for treating or preventing a neuronal disorder (e.g. akathesia,
CC Alzheimer's disease, amnesia, dementia or parkinson's disease),
CC connective tissue disorder (e.g. cystic fibrosis, myocardial fibrosis,
CC osteoporosis or Marfan syndrome), or a disorder of cell proliferation
CC (e.g. cirrhosis, leukaemia, lymphoma, sarcoma or bone marrow cancer).
CC The polypeptide or its fragments are also useful for screening
CC libraries of compounds in any of several drug-screening techniques.

XX
Sequence 314 AA;

Query Match 100.0%; Score 1567; DB 21; Length 314;

Best Local Similarity 100.0%; Pred. No. 1.9e-159;

Matches 314; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSVGRIGAGOLAYRFTAAIGILSAHKITASPRENLPTVSALKRKGYNLTRSNKETVKHD 60
Db 1 msyfgifagqlayrftaaigilshakiiaspennlptvsalkrkgynltsnketykhsd 60
QY 61 VLFALKPHIFILDEGADYQARHIVYSCAGVTISVERKLMQAFQAPKVRIMNT 120
Db 61 vlfalkphifildegadqarhivscaagvtisvekklmqafqapkvircnt 120
QY 121 PVVDEGATWYATGTHALVEDGQLEQMLSSVGFCTEVEEDIDAVYGLSESQPAYAFMA 180
Db 121 mtntpvvvqegatwyatgthalvedgqleqmlssvgfcfeeedidavtglsqspay 180
Db 121 pvvqegatwyatgthalvedgqleqmlssvgfcfeeedidavtglsqspayfma 180
QY 181 LDADGEVKMGLPRTAQLQGAQALIGAKMILDSSEQHQCQKVNCSPPGATHALHFLE 240
Db 181 afdmaladagvgkingplrraqlqgaaqlgaakmlldseqhpcqklkdnvcspgatihalfle 240
QY 241 SGGRSILINNEAVASCIRTRLQSMQEKISPAALKTLLDRVKLESPTVSPWLTSPSG 300
Db 241 sggfrsllineavascirtrlqsmadekispaalktlldrvklesptvstltpssp 300
QY 301 KULTRSLAIGRKD 314
Db 301 kultrsaligrkd 314

XX
Sequence 320 AA;

Query Match 96.3%; Score 1509; DB 22; Length 320;

Best Local Similarity 96.9%; Pred. No. 3.3e-153;

Matches 310; Conservative 0; Mismatches 4; Indels 6; Gaps 2; Gaps 2;

QY 1 MSVFIGQOLAYR---FTAGILSAHKITASPRENLPTVSALKRKGYNLTRSNKETV 56
Db 1 msyfgifagqlanalarfgtaagilshakiiaspennlptvsalkrkgynltsnkety 60
QY 57 KHSDFVLFLAKPHIFILDEGADYQARHIVYSCAGVTISVERKLMQAFQAPKVRIC 116
Db 61 khsdfvlakphifildegadqarhivscaagvtisvekklmqafqapkvirc 116
QY 117 MNTNPVWVQEGATWYATGTHALVEDGQLEQMLSSVGFCTEVEEDIDAVYGLSGSPAY 176
Db 121 mtntpvvvqegatwyatgthalvedgqleqmlssvgfcfeeedidavtglsqspay 180
Db 177 AFMAID--DGGYVKMGLPRTAQLQGAQALIGAKMILDSSEQHQCQKVNCSPPGATH 234
QY 181 afdmaladagvgkingplrraqlqgaaqlgaakmlldseqhpcqklkdnvcspgatihalfle 240
QY 235 ALHELESGGRSILINNEAVASCIRTRLQSMQEKISPAALKTLLDRVKLESPTVSPWLTSPSG 294
Db 241 alhiflesgrsllineavascirtrlqsmadekispaalktlldrvklesptvstl 300
QY 295 TPSSGKLTSLAIGKKD 314
Db 301 tpspgklttsaligkkd 320

XX
Sequence 319 AA;

Query Match 96.3%; Score 1509; DB 22; Length 319 AA;

Best Local Similarity 96.9%; Pred. No. 3.3e-153;

Matches 310; Conservative 0; Mismatches 4; Indels 6; Gaps 2; Gaps 2;

QY 1 MSVFIGQOLAYR---FTAGILSAHKITASPRENLPTVSALKRKGYNLTRSNKETV 56
AC 57 KHSDFVLFLAKPHIFILDEGADYQARHIVYSCAGVTISVERKLMQAFQAPKVRIC 116
XX
AC 121 mtntpvvvqegatwyatgthalvedgqleqmlssvgfcfeeedidavtglsqspay 180
XX
AC 177 AFMAID--DGGYVKMGLPRTAQLQGAQALIGAKMILDSSEQHQCQKVNCSPPGATH 234
XX
AC 181 afdmaladagvgkingplrraqlqgaaqlgaakmlldseqhpcqklkdnvcspgatihalfle 240
XX
AC 235 ALHELESGGRSILINNEAVASCIRTRLQSMQEKISPAALKTLLDRVKLESPTVSPWLTSPSG 294
XX
AC 241 alhiflesgrsllineavascirtrlqsmadekispaalktlldrvklesptvstl 300
XX
AC 295 TPSSGKLTSLAIGKKD 314
XX
AC 301 tpspgklttsaligkkd 320

XX
Sequence 319 AA;

Query Match 96.3%; Score 1509; DB 22; Length 319 AA;

Best Local Similarity 96.9%; Pred. No. 3.3e-153;

Matches 310; Conservative 0; Mismatches 4; Indels 6; Gaps 2; Gaps 2;

QY 1 MSVFIGQOLAYR---FTAGILSAHKITASPRENLPTVSALKRKGYNLTRSNKETV 56
AC 121 mtntpvvvqegatwyatgthalvedgqleqmlssvgfcfeeedidavtglsqspay 180
XX
AC 177 AFMAID--DGGYVKMGLPRTAQLQGAQALIGAKMILDSSEQHQCQKVNCSPPGATH 234
XX
AC 181 afdmaladagvgkingplrraqlqgaaqlgaakmlldseqhpcqklkdnvcspgatihalfle 240
XX
AC 235 ALHELESGGRSILINNEAVASCIRTRLQSMQEKISPAALKTLLDRVKLESPTVSPWLTSPSG 294
XX
AC 241 alhiflesgrsllineavascirtrlqsmadekispaalktlldrvklesptvstl 300
XX
AC 295 TPSSGKLTSLAIGKKD 314
XX
AC 301 tpspgklttsaligkkd 320

XX
Sequence 319 AA;

Query Match 96.3%; Score 1509; DB 22; Length 319 AA;

Best Local Similarity 96.9%; Pred. No. 3.3e-153;

Matches 310; Conservative 0; Mismatches 4; Indels 6; Gaps 2; Gaps 2;

QY 1 MSVFIGQOLAYR---FTAGILSAHKITASPRENLPTVSALKRKGYNLTRSNKETV 56
AC 121 mtntpvvvqegatwyatgthalvedgqleqmlssvgfcfeeedidavtglsqspay 180
XX
AC 177 AFMAID--DGGYVKMGLPRTAQLQGAQALIGAKMILDSSEQHQCQKVNCSPPGATH 234
XX
AC 181 afdmaladagvgkingplrraqlqgaaqlgaakmlldseqhpcqklkdnvcspgatihalfle 240
XX
AC 235 ALHELESGGRSILINNEAVASCIRTRLQSMQEKISPAALKTLLDRVKLESPTVSPWLTSPSG 294
XX
AC 241 alhiflesgrsllineavascirtrlqsmadekispaalktlldrvklesptvstl 300
XX
AC 295 TPSSGKLTSLAIGKKD 314
XX
AC 301 tpspgklttsaligkkd 320

XX
Sequence 319 AA;

Query Match 96.3%; Score 1509; DB 22; Length 319 AA;

Best Local Similarity 96.9%; Pred. No. 3.3e-153;

Matches 310; Conservative 0; Mismatches 4; Indels 6; Gaps 2; Gaps 2;

QY 1 MSVFIGQOLAYR---FTAGILSAHKITASPRENLPTVSALKRKGYNLTRSNKETV 56
AC 121 mtntpvvvqegatwyatgthalvedgqleqmlssvgfcfeeedidavtglsqspay 180
XX
AC 177 AFMAID--DGGYVKMGLPRTAQLQGAQALIGAKMILDSSEQHQCQKVNCSPPGATH 234
XX
AC 181 afdmaladagvgkingplrraqlqgaaqlgaakmlldseqhpcqklkdnvcspgatihalfle 240
XX
AC 235 ALHELESGGRSILINNEAVASCIRTRLQSMQEKISPAALKTLLDRVKLESPTVSPWLTSPSG 294
XX
AC 241 alhiflesgrsllineavascirtrlqsmadekispaalktlldrvklesptvstl 300
XX
AC 295 TPSSGKLTSLAIGKKD 314
XX
AC 301 tpspgklttsaligkkd 320

XX
Sequence 319 AA;

Query Match 96.3%; Score 1509; DB 22; Length 319 AA;

Best Local Similarity 96.9%; Pred. No. 3.3e-153;

Matches 310; Conservative 0; Mismatches 4; Indels 6; Gaps 2; Gaps 2;

QY 1 MSVFIGQOLAYR---FTAGILSAHKITASPRENLPTVSALKRKGYNLTRSNKETV 56
AC 121 mtntpvvvqegatwyatgthalvedgqleqmlssvgfcfeeedidavtglsqspay 180
XX
AC 177 AFMAID--DGGYVKMGLPRTAQLQGAQALIGAKMILDSSEQHQCQKVNCSPPGATH 234
XX
AC 181 afdmaladagvgkingplrraqlqgaaqlgaakmlldseqhpcqklkdnvcspgatihalfle 240
XX
AC 235 ALHELESGGRSILINNEAVASCIRTRLQSMQEKISPAALKTLLDRVKLESPTVSPWLTSPSG 294
XX
AC 241 alhiflesgrsllineavascirtrlqsmadekispaalktlldrvklesptvstl 300
XX
AC 295 TPSSGKLTSLAIGKKD 314
XX
AC 301 tpspgklttsaligkkd 320

XX
Sequence 319 AA;

Query Match 96.3%; Score 1509; DB 22; Length 319 AA;

Best Local Similarity 96.9%; Pred. No. 3.3e-153;

Matches 310; Conservative 0; Mismatches 4; Indels 6; Gaps 2; Gaps 2;

QY 1 MSVFIGQOLAYR---FTAGILSAHKITASPRENLPTVSALKRKGYNLTRSNKETV 56
AC 121 mtntpvvvqegatwyatgthalvedgqleqmlssvgfcfeeedidavtglsqspay 180
XX
AC 177 AFMAID--DGGYVKMGLPRTAQLQGAQALIGAKMILDSSEQHQCQKVNCSPPGATH 234
XX
AC 181 afdmaladagvgkingplrraqlqgaaqlgaakmlldseqhpcqklkdnvcspgatihalfle 240
XX
AC 235 ALHELESGGRSILINNEAVASCIRTRLQSMQEKISPAALKTLLDRVKLESPTVSPWLTSPSG 294
XX
AC 241 alhiflesgrsllineavascirtrlqsmadekispaalktlldrvklesptvstl 300
XX
AC 295 TPSSGKLTSLAIGKKD 314
XX
AC 301 tpspgklttsaligkkd 320

XX
Sequence 319 AA;

Query Match 96.3%; Score 1509; DB 22; Length 319 AA;

Best Local Similarity 96.9%; Pred. No. 3.3e-153;

Matches 310; Conservative 0; Mismatches 4; Indels 6; Gaps 2; Gaps 2;

QY 1 MSVFIGQOLAYR---FTAGILSAHKITASPRENLPTVSALKRKGYNLTRSNKETV 56
AC 121 mtntpvvvqegatwyatgthalvedgqleqmlssvgfcfeeedidavtglsqspay 180
XX
AC 177 AFMAID--DGGYVKMGLPRTAQLQGAQALIGAKMILDSSEQHQCQKVNCSPPGATH 234
XX
AC 181 afdmaladagvgkingplrraqlqgaaqlgaakmlldseqhpcqklkdnvcspgatihalfle 240
XX
AC 235 ALHELESGGRSILINNEAVASCIRTRLQSMQEKISPAALKTLLDRVKLESPTVSPWLTSPSG 294
XX
AC 241 alhiflesgrsllineavascirtrlqsmadekispaalktlldrvklesptvstl 300
XX
AC 295 TPSSGKLTSLAIGKKD 314
XX
AC 301 tpspgklttsaligkkd 320

XX
Sequence 319 AA;

Query Match 96.3%; Score 1509; DB 22; Length 319 AA;

Best Local Similarity 96.9%; Pred. No. 3.3e-153;

Matches 310; Conservative 0; Mismatches 4; Indels 6; Gaps 2; Gaps 2;

QY 1 MSVFIGQOLAYR---FTAGILSAHKITASPRENLPTVSALKRKGYNLTRSNKETV 56
AC 121 mtntpvvvqegatwyatgthalvedgqleqmlssvgfcfeeedidavtglsqspay 180
XX
AC 177 AFMAID--DGGYVKMGLPRTAQLQGAQALIGAKMILDSSEQHQCQKVNCSPPGATH 234
XX
AC 181 afdmaladagvgkingplrraqlqgaaqlgaakmlldseqhpcqklkdnvcspgatihalfle 240
XX
AC 235 ALHELESGGRSILINNEAVASCIRTRLQSMQEKISPAALKTLLDRVKLESPTVSPWLTSPSG 294
XX
AC 241 alhiflesgrsllineavascirtrlqsmadekispaalktlldrvklesptvstl 300
XX
AC 295 TPSSGKLTSLAIGKKD 314
XX
AC 301 tpspgklttsaligkkd 320

XX
Sequence 319 AA;

Query Match 96.3%; Score 1509; DB 22; Length 319 AA;

Best Local Similarity 96.9%; Pred. No. 3.3e-153;

Matches 310; Conservative 0; Mismatches 4; Indels 6; Gaps 2; Gaps 2;

QY 1 MSVFIGQOLAYR---FTAGILSAHKITASPRENLPTVSALKRKGYNLTRSNKETV 56
AC 121 mtntpvvvqegatwyatgthalvedgqleqmlssvgfcfeeedidavtglsqspay 180
XX
AC 177 AFMAID--DGGYVKMGLPRTAQLQGAQALIGAKMILDSSEQHQCQKVNCSPPGATH 234
XX
AC 181 afdmaladagvgkingplrraqlqgaaqlgaakmlldseqhpcqklkdnvcspgatihalfle 240
XX
AC 235 ALHELESGGRSILINNEAVASCIRTRLQSMQEKISPAALKTLLDRVKLESPTVSPWLTSPSG 294
XX
AC 241 alhiflesgrsllineavascirtrlqsmadekispaalktlldrvklesptvstl 300
XX
AC 295 TPSSGKLTSLAIGKKD 314
XX
AC 301 tpspgklttsaligkkd 320

XX
Sequence 319 AA;

Query Match 96.3%; Score 1509; DB 22; Length 319 AA;

Best Local Similarity 96.9%; Pred. No. 3.3e-153;

Matches 310; Conservative 0; Mismatches 4; Indels 6; Gaps 2; Gaps 2;

QY 1 MSVFIGQOLAYR---FTAGILSAHKITASPRENLPTVSALKRKGYNLTRSNKETV 56
AC 121 mtntpvvvqegatwyatgthalvedgqleqmlssvgfcfeeedidavtglsqspay 180
XX
AC 177 AFMAID--DGGYVKMGLPRTAQLQGAQALIGAKMILDSSEQHQCQKVNCSPPGATH 234
XX
AC 181 afdmaladagvgkingplrraqlqgaaqlgaakmlldseqhpcqklkdnvcspgatihalfle 240
XX
AC 235 ALHELESGGRSILINNEAVASCIRTRLQSMQEKISPAALKTLLDRVKLESPTVSPWLTSPSG 294
XX
AC 241 alhiflesgrsllineavascirtrlqsmadekispaalktlldrvklesptvstl 300
XX
AC 295 TPSSGKLTSLAIGKKD 314
XX
AC 301 tpspgklttsaligkkd 320

XX
Sequence 319 AA;

Query Match 96.3%; Score 1509; DB 22; Length 319 AA;

Best Local Similarity 96.9%; Pred. No. 3.3e-153;

Matches 310; Conservative 0; Mismatches 4; Indels 6; Gaps 2; Gaps 2;

QY 1 MSVFIGQOLAYR---FTAGILSAHKITASPRENLPTVSALKRKGYNLTRSNKETV 56
AC 121 mtntpvvvqegatwyatgthalvedgqleqmlssvgfcfeeedidavtglsqspay 180
XX
AC 177 AFMAID--DGGYVKMGLPRTAQLQGAQALIGAKMILDSSEQHQCQKVNCSPPGATH 234
XX
AC 181 afdmaladagvgkingplrraqlqgaaqlgaakmlldseqhpcqklkdnvcspgatihalfle 240
XX
AC 235 ALHELESGGRSILINNEAVASCIRTRLQSMQEKISPAALKTLLDRVKLESPTVSPWLTSPSG 294
XX
AC 241 alhiflesgrsllineavascirtrlqsmadekispaalktlldrvklesptvstl 300
XX
AC 295 TPSSGKLTSLAIGKKD 314
XX
AC 301 tpspgklttsaligkkd 320

XX
Sequence 319 AA;

Query Match 96.3%; Score 1509; DB 22; Length 319 AA;

Best Local Similarity 96.9%; Pred. No. 3.3e-153;

Matches 310; Conservative 0; Mismatches 4; Indels 6; Gaps 2; Gaps 2;

QY 1 MSVFIGQOLAYR---FTAGILSAHKITASPRENLPTVSALKRKGYNLTRSNKETV 56
AC 121 mtntpvvvqegatwyatgthalvedgqleqmlssvgfcfeeedidavtglsqspay 180
XX
AC 177 AFMAID--DGGYVKMGLPRTAQLQGAQALIGAKMILDSSEQHQCQKVNCSPPGATH 234
XX
AC 181 afdmaladagvgkingplrraqlqgaaqlgaakmlldseqhpcqklkdnvcspgatihalfle 240
XX
AC 235 ALHELESGGRSILINNEAVASCIRTRLQSMQEKISPAALKTLLDRVKLESPTVSPWLTSPSG 294
XX
AC 241 alhiflesgrsllineavascirtrlqsmadekispaalktlldrvklesptvstl 300
XX
AC 295 TPSSGKLTSLAIGKKD 314
XX
AC 301 tpspgklttsaligkkd 320

XX
Sequence 319 AA;

Query Match 96.3%; Score 1509; DB 22; Length 319 AA;

Best Local Similarity 96.9%; Pred. No. 3.3e-153;

Matches 310; Conservative 0; Mismatches 4; Indels 6; Gaps 2; Gaps 2;

QY 1 MSVFIGQOLAYR---FTAGILSAHKITASPRENLPTVSALKRKGYNLTRSNKETV 56
AC 121 mtntpvvvqegatwyatgthalvedgqleqmlssvgfcfeeedidavtglsqspay 180
XX
AC 177 AFMAID--DGGYVKMGLPRTAQLQGAQALIGAKMILDSSEQHQCQKVNCSPPGATH 234
XX
AC 181 afdmaladagvgkingplrraqlqgaaqlgaakmlldseqhpcqklkdnvcspgatihalfle 240
XX
AC 235 ALHELESGGRSILINNEAVASCIRTRLQSMQEKISPAALKTLLDRVKLESPTVSPWLTSPSG 294
XX
AC 241 alhiflesgrsllineavascirtrlqsmadekispaalktlldrvklesptvstl 300
XX
AC 295 TPSSGKLTSLAIGKKD 314
XX
AC 301 tpspgklttsaligkkd 320

XX
Sequence 319 AA;

Query Match 96.3%; Score 1509; DB 22; Length 319 AA;

Best Local Similarity 96.9%; Pred. No. 3.3e-153;

Matches 310; Conservative 0; Mismatches 4; Indels 6; Gaps 2; Gaps 2;

QY 1 MSVFIGQOLAYR---FTAGILSAHKITASPRENLPTVSALKRKGYNLTRSNKETV 56
AC 121 mtntpvvvqegatwyatgthalvedgqleqmlssvgfcfeeedidavtglsqspay 180
XX
AC 177 AFMAID--DGGYVKMGLPRTAQLQGAQALIGAKMILDSSEQHQCQKVNCSPPGATH 234
XX
AC 181 afdmaladagvgkingplrraqlqgaaqlgaakmlldseqhpcqklkdnvcspgatihalfle 240
XX
AC 235 ALHELESGGRSILINNEAVASCIRTRLQSMQEKISPAALKTLLDRVKLESPTVSPWLTSPSG 294
XX
AC 241 alhiflesgrsllineavascirtrlqsmadekispaalktlldrvklesptvstl 300
XX
AC 295 TPSSGKLTSLAIGKKD 314
XX
AC 301 tpspgklttsaligkkd 320

XX
Sequence 319 AA;

Query Match 96.3%; Score 1509; DB 22; Length 319 AA;

Best Local Similarity 96.9%; Pred. No. 3.3e-153;

Matches 310; Conservative 0; Mismatches 4; Indels 6; Gaps 2; Gaps 2;

QY 1 MSVFIGQOLAYR---FTAGILSAHKITASPRENLPTVSALKRKGYNLTRSNKETV 56
AC 121 mtntpvvvqegatwyatgthalvedgqleqmlssvgfcfeeedidavtglsqspay 180
XX
AC 177 AFMAID--DGGYVKMGLPRTAQLQGAQALIGAKMILDSSEQHQCQKVNCSPPGATH 234
XX
AC 181 afdmaladagvgkingplrraqlqgaaqlgaakmlldseqhpcqklkdnvcspgatihalfle 240
XX
AC 235 ALHELESGGRSILINNEAVASCIRTRLQSMQEKISPAALKTLLDRVKLESPTVSPWLTSPSG 294
XX
AC 241 alhiflesgrsllineavascirtrlqsmadekispaalktlldrvklesptvstl 300
XX
AC 295 TPSSGKLTSLAIGKKD 314
XX
AC 301 tpspgklttsaligkkd 320

XX
Sequence 319 AA;

Query Match 96.3%; Score 1509; DB 22; Length 319 AA;

Best Local Similarity 96.9%; Pred. No. 3.3e-153;

Matches 310; Conservative 0; Mismatches 4; Indels 6; Gaps 2; Gaps 2;

QY 1 MSVFIGQOLAYR---FTAGILSAHKITASPRENLPTVSALKRKGYNLTRSNKETV 56
AC 121 mtntpvvvqegatwyatgthalvedgqleqmlssvgfcfeeedidavtglsqspay 180
XX
AC 177 AFMAID--DGGYVKMGLPRTAQLQGAQALIGAKMILDSSEQHQCQKVNCSPPGATH 234
XX
AC 181 afdmaladagvgkingplrraqlqgaaqlgaakmlldseqhpcqklkdnvcspgatihalfle 240
XX
AC 235 ALHELESGGRSILINNEAVASCIRTRLQSMQEKISPAALKTLLDRVKLESPTVSPWLTSPSG 294
XX
AC 241 alhiflesgrsllineavascirtrlqsmadekispaalktlldrvklesptvstl 300
XX
AC 295 TPSSGKLTSLAIGKKD 314
XX
AC 301 tpspgklttsaligkkd 320

XX
Sequence 319 AA;

Query Match 96.3%; Score 1509; DB 22; Length 319 AA;

Best Local Similarity 96.9%; Pred. No. 3.3e-153;

Matches 310; Conservative 0; Mismatches 4; Indels 6; Gaps 2; Gaps 2;

QY 1 MSVFIGQ

OS Homo sapiens.
 XX
 PN CN1274728-A.
 XX
 PD 29-NOV-2000.
 XX
 PR 25-MAY-1999; 99CN-0107071.
 XX
 PA (UYFU-) UNIV FUDAN.
 XX
 PT Yu L, Fu Q, Zhang H;
 XX
 DR WPI: 2001-211749/2.
 XX
 PT New human protein and its code sequence, preparation and application -
 PS Example 2; Fig 1; 20pp; Chinese.
 XX
 CC The present invention describes a human protein designated Py-CR, which
 CC is homologous to the human pyrrole-5'-carboxylate reductase (P5CR)
 CC (EC1.5.1.2). The present invention also describes methods for the
 CC application and production process of the Py-CR polynucleotide and
 CC protein which is used in comparison with the human P5CR
 CC an example from the present invention.
 XX
 SQ Sequence 319 AA;

Query Match 83.3%; Score 1306; DB 22; Length 319;
 Best Local Similarity 82.7%; Pred. No. 1.9e-131;
 Matches 263; Conservative 27; Mismatches 22; Indels 6; Gaps 2;
 Oy 1 MSVPTGACOLAVR---FTAGILSAHKIASPPENNLPTPSALRKGVNLTRSNKETV 56
 Db 1 msVptgacolavr---ftagilshkiasppennlptpsalrkgnltv 60
 Oy 57 KHSVLFVAKPPTIPFLDEGADVQARHIVSCAGVTISSVERKLMFOPAKVIRC 116
 Db 61 qhadvlflavkphipflfdegaedrlhivscaagtisskkksafrpaprvc 120
 Oy 117 MTPVPPVWQEGATVYATGTHALVDEQOLGQOLMSVGFCVEEVDLIDAVTGLSGSGPAY 176
 Db 121 mtptpvvregatvyalghaqedrlhivscaagtisskkksafrpaprvc 180
 Oy 177 AFWALD---ADGGVKMGLPRLAIAQGAQALIGAAMLLSEQHQPCQLKONVCSRGATH 234
 Db 181 afraaldaladsgvkmglprlavravrgaqaligaakmllhseqhpgqkdnvsspgatih 240
 Oy 235 ATRPLEGGFRSLLNAVEASCIRRELOQSMADQEKISPAALKTLIDRVKLBSPTVSTL 294
 Db 241 alhvleesggffslinavaeascirrtreloqmadqeqvpaakktildkvkldspagtal 300
 Oy 295 TPSSPKLTLTSLAAGK 312
 Db 301 spsghtklprslapagk 318
 RESULT 4
 AAB93027 standard; Protein; 212 AA.
 ID AAB93027
 AC
 XX
 DT 26-JUN-2001 (first entry)
 DE Human protein sequence SEQ ID NO:11795.
 KW Human; primer; detection; diagnosis; antisense therapy; gene therapy.
 OS Homo sapiens.

XX
 PN EP1074617-A2.
 XX
 PD 07-FEB-2001.
 XX
 PR 28-JUL-2000; 2000EP-0116126.
 XX
 PR 29-JUL-1999; 99JP-0248036.
 PR 21-AUG-1999; 99JP-030253.
 PR 11-JAN-2000; 2000JP-011376.
 PR 02-MAY-2000; 2000JP-0183767.
 PR 09-JUN-2000; 2000JP-0241899.
 XX
 PA (HELI-) HELIX RES INST.
 XX
 PI Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
 PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
 XX
 DR WPI; 2001-318749/34.
 XX
 PR Primer sets for synthesizing polynucleotides, particularly the 5602
 PR full-length cDNAs defined in the specification, and for the detection
 PR and/or diagnosis of the abnormality of the proteins encoded by the
 PR full-length cDNAs -
 XX
 PS Claim 8; SEQ ID 11795; 2537pp + CD ROM; English.
 XX
 CC The present invention describes primer sets for synthesising 5602
 CC full-length cDNAs defined in the specification. Where a primer set
 CC comprises: (a) an oligo-dT primer and an oligonucleotide complementary
 CC to the complementary strand of a polynucleotide which comprises one of
 CC the 5602 nucleotide sequences defined in the specification, where the
 CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
 CC of an oligonucleotide comprising a sequence complementary to the
 CC complementary strand of a polynucleotide which comprises a 5'-end
 CC polynucleotide which comprises a 3'-end sequence, where the
 CC oligonucleotide comprises at least 15 nucleotides and the combination of
 CC the 5'-end sequence/3'-end sequence is selected from those defined in
 CC the specification. The primer sets can be used in antisense therapy and
 CC in gene therapy. The primers are useful for synthesising polynucleotides,
 CC particularly, full-length cDNAs. The primers are also useful for the
 CC detection and/or diagnosis of the abnormality of the proteins encoded by
 CC the full-length cDNAs. The primers allow obtaining of the full-length
 CC cDNAs easily without any specialised methods. AAB03166 to AAH13628 and
 CC AAH13633 to AAH18742 represent human cDNA sequences; AAB02446 to
 CC AAB95893 represent human amino acid sequences; and AAB13629 to AAH13632
 CC represent oligonucleotides, all of which are used in the exemplification
 XX
 SQ Sequence 212 AA;
 Query Match 66.7%; Score 1045; DB 22; Length 212;
 Best Local Similarity 99.1%; Pred. No. 1e-103;
 Matches 210; Conservative 0; Mismatches 0; Indels 2; Gaps 1;
 Oy 105 MAFOPAKVKIRCMTPTVWQEGATVYATGTHALVDEQOLGQOLMSVGFCVEEVDLID 164
 Db 1 matqpkvirkmtptvvwqegatvtyatgthalvdeqoleqimsvgfcveedlid 60
 Oy 165 AVTGLSGSPAYAFWALD---ADGGVKMGLPRLAIAQGAQALIGAAMLLSEQHQPCQLK 222
 Db 61 avtglsqsgayafmaladsgvkmglprlavravrgaqaligaakmllhseqhpgqkdnvsspgatih 120
 Oy 223 DNWSPGGATIHAEFLLEGGFRSLLNAVEASCIRRELOQSMADQEKISPAALKTLID 282
 Db 121 dnwspggatihfllesggfrslinavaeascirrtreloqmadqeqvpaakktildkvkldspagtal
 Oy 283 RVKESPTVSTLTPSSPGKLLTSLAAGKKD 314
 Db 181 rvklesptvstltpsspgkllttslaagkkd 212

RESULT 5

ID AAB56640 standard: Protein; 255 AA.

XX AAB56640;

XX DT 13-MAR-2001 (first entry)

XX DE Human prostate cancer antigen protein sequence SEQ ID NO:1218.

XX KW Human; prostate cancer; prostate cancer antigen; detection; diagnosis; neuroprotective; cytostatic; cardiotrostatic; immunomodulatory; muscular; vulnerable; gastrointestinal; nephroprotective; antinfective; gynaecological; antibacterial; gene therapy; neural; immune; reproductive; renal; gastrointestinal; pulmonary; cardiovascular; proliferative; disorder; wound; infectious disease.

XX OS Homo sapiens.

XX PN WO200055174-A1.

XX PD 21-SEP-2000.

XX PE 08-MAR-2000; 20000WO-US05988.

XX PR 12-MAR-1999; 99US-0124270.

XX PA (HUMA-) HUMAN GENOME SCI INC.

XX PA (ROSE/) ROSEN C A.

XX PT Rosen CA, Ruben SM;

XX DR WPI: 2000-587513-55.

XX DR N-PSDB; AAF15843.

XX XAF15566 to AAF16505 encode the human prostate cancer associated proteins, called prostate cancer antigens, given in AAB56363 to AAB57302. The prostate cancer antigens can have neuroprotective, cytostatic, cardioactive, immunomodulatory, muscular, vulnerable, gastrointestinal, nephroprotective, antinfective, gynaecological and antibacterial activities, and can be used in gene therapy. The prostate cancer antigen polynucleotides may be used for detection of prostate cancer, chromosome identification, as chromosome markers, and for numerous other diagnostic or research purposes. The prostate cancer antigens may be used to treat disorders such as neural, immune, muscular, reproductive, gastrointestinal, pulmonary, cardiovascular, renal, and proliferative disorders, wounds, and infectious diseases. AAF16506 to AAF16514 to CC AAB57303 represent sequences used in the exemplification of the present invention.

XX Sequence 255 AA;

Query Match 66.5%; Score 1042; DB 21; Length 255; Best Local Similarity 84.6%; Pred. No. 2.8e-103; Matches 208; Conservative 19; Mismatches 17; Indels 2; Gaps 1; Matches 208; Conservative 19; Mismatches 17; Indels 2; Gaps 1;

QY 69 HIIPTFIDEGIADQARHIVVSCAAGVTISSEKIMAFQPAKVRQMTNPVWYQEGA 128

Db 9 hiiptfidegiadqarhivvscagvtissekimafqpakvrqmtnpvrega 68

Db 9 hiiptfidegiadqarhivvscagvtissekimafqpakvrqmtnpvrega 68

QY 129 TTYATGTHALEDGQLEQMLMSVQPCTEWEEDLDAVGLSGSGPAYAFMADL-ADCG 186

Db 69 tyyatgthaqedgqleqmlmsvqpcfteedladvglsgsgpayaftaldaladgg 128

QY 187 VRMGLPPLATQIGAQLGAKMLDSEPRCQLKDNVCSPGGATTHALHLESGGRS 246

RESULT 6

ID AAY92517 standard: Protein; 274 AA.

XX AAY92517;

XX AC AAY92517;

XX DT 10-AUG-2000 (first entry)

XX DE Human OXRE-14.

XX KW OXRE-14; oxidoreductase; pyrrole-5-carboxylate-reductase; anti-tumor; neurotropic; neuroprotectant; antiparkinsonian; S, antisclerotic; anxiolytic; antischizophrenic; anti-inflammatory; antiarthritic.

XX OS Homo sapiens.

XX PA Key

FT Location/Qualifiers

FT Region 9.256

FT /label= "signature-sequence

FT /note= "for delta 1-pyrrole-5-carboxylate reductase"

FT Modified-site 7

FT /note= "Potential phosphorylation site"

FT Modified-site 44

FT /note= "potential Phosphorylation site"

FT Modified-site 61

FT /note= "potential phosphorylation site"

FT Modified-site 108

FT /note= "potential phosphorylation site"

FT Modified-site 109

FT /note= "potential phosphorylation site"

FT Modified-site 216

FT /note= "potential phosphorylation site"

FT Modified-site 259

FT /note= "potential phosphorylation site"

FT Modified-site 265

FT /note= "potential phosphorylation site"

XX PN WO200020604-A2.

XX PD 13-APR-2000.

XX PR 06-OCT-1999; 99WO-US23434.

XX PR 06-OCT-1998; 98US-0155202.

XX PR 02-DEC-1998; 99US-0123911.

XX PR 10-MAR-1999;

XX PA (INCY-) INCYTE PHARM INC.

XX PI Lal P, Guegler KJ, Corley NC, Baugn MR, Tang YT, Lu DAM;

XX PI Hillman JL, Bandman O, Azimzai Y, Au-Young J, Yue H;

XX PI Yang J;

XX DR WPI: 2000-303785/26.

XX DR N-PSDB; AIA09388.

XX PT Purified polypeptide for treating or preventing disorders associated with decreased expression or activity of oxidoreductase molecules

XX Claim 1; Page 84; 97pp; English.

Arabidopsis thaliana.		Protein identification; signal hybridisation assay; genetic map terminant sequence.
X		EP103405-R2.
X		06-SEP-2000.
D		25-FEB-2000; 2000EP-0301439.
X		25-FEB-1999; 99US-0121825.
X		05-MAR-1999; 99US-0123180.
X		09-MAR-1999; 99US-0123548.
X		23-MAR-1999; 99US-0125788.
X		25-MAR-1999; 99US-0126264.
X		29-MAR-1999; 99US-0126785.
X		01-APR-1999; 99US-0127462.
X		06-APR-1999; 99US-0128234.
X		08-APR-1999; 99US-0128714.
X		16-APR-1999; 99US-0129845.
X		19-APR-1999; 99US-0130077.
X		21-APR-1999; 99US-0130049.
X		23-APR-1999; 99US-0130510.
X		23-APR-1999; 99US-0130891.
X		28-APR-1999; 99US-0131449.
X		30-APR-1999; 99US-0132048.
X		30-APR-1999; 99US-0132407.
X		04-MAY-1999; 99US-0132484.
X		05-MAY-1999; 99US-0132485.
X		06-MAY-1999; 99US-0132486.
X		06-MAY-1999; 99US-0132487.
X		07-MAY-1999; 99US-0132488.
X		11-MAY-1999; 99US-0134266.
X		14-MAY-1999; 99US-0134284.
X		14-MAY-1999; 99US-0134285.
X		14-MAY-1999; 99US-0134286.
X		14-MAY-1999; 99US-0134287.
X		14-MAY-1999; 99US-0134288.
X		18-MAY-1999; 99US-0134289.
X		19-MAY-1999; 99US-0134941.
X		20-MAY-1999; 99US-0135124.
X		21-MAY-1999; 99US-0135533.
X		24-MAY-1999; 99US-0135629.
X		25-MAY-1999; 99US-0136021.
X		27-MAY-1999; 99US-0136392.
X		28-MAY-1999; 99US-0136782.
X		01-JUN-1999; 99US-0137222.
X		03-JUN-1999; 99US-0137528.
X		04-JUN-1999; 99US-0137500.
X		07-JUN-1999; 99US-0137724.
X		08-JUN-1999; 99US-0138094.
X		10-JUN-1999; 99US-0138592.
X		10-JUN-1999; 99US-0138847.
X		14-JUN-1999; 99US-0139452.
X		16-JUN-1999; 99US-0139455.
X		16-JUN-1999; 99US-0139458.
X		18-JUN-1999; 99US-0139453.
X		18-JUN-1999; 99US-0139460.
X		18-JUN-1999; 99US-0139461.
X		18-JUN-1999; 99US-0139454.
X		18-JUN-1999; 99US-0139463.
X		18-JUN-1999; 99US-0139457.
X		18-JUN-1999; 99US-0139750.
X		18-JUN-1999; 99US-0139763.
X		21-JUN-1999; 99US-0139877.
X		22-JUN-1999; 99US-0139899.
X		23-JUN-1999; 99US-0140353.

PR 23-JUN-1999; 99US-0140354.
 PR 24-JUN-1999; 99US-0140695.
 PR 28-JUN-1999; 99US-0140823.
 PR 29-JUN-1999; 99US-0140991.
 PR 30-JUN-1999; 99US-0141287.
 PR 01-JUL-1999; 99US-0141842.
 PR 01-JUL-1999; 99US-0142154.
 PR 02-JUL-1999; 99US-0142055.
 PR 06-JUL-1999; 99US-0142390.
 PR 08-JUL-1999; 99US-0142803.
 PR 09-JUL-1999; 99US-0142920.
 PR 12-JUL-1999; 99US-0142977.
 PR 13-JUL-1999; 99US-0143542.
 PR 14-JUL-1999; 99US-0143624.
 PR 15-JUL-1999; 99US-0144005.
 PR 16-JUL-1999; 99US-0144085.
 PR 16-JUL-1999; 99US-0144086.
 PR 19-JUL-1999; 99US-0144325.
 PR 19-JUL-1999; 99US-0144331.
 PR 19-JUL-1999; 99US-0144332.
 PR 19-JUL-1999; 99US-0144333.
 PR 19-JUL-1999; 99US-0144334.
 PR 19-JUL-1999; 99US-0144335.
 PR 20-JUL-1999; 99US-0144352.
 PR 20-JUL-1999; 99US-0144632.
 PR 22-JUL-1999; 99US-0144884.
 PR 21-JUL-1999; 99US-0144814.
 PR 23-JUL-1999; 99US-0145086.
 PR 21-JUL-1999; 99US-0145088.
 PR 22-JUL-1999; 99US-0145085.
 PR 22-JUL-1999; 99US-0145087.
 PR 22-JUL-1999; 99US-0145089.
 PR 22-JUL-1999; 99US-0145192.
 PR 23-JUL-1999; 99US-0145145.
 PR 23-JUL-1999; 99US-0145218.
 PR 23-JUL-1999; 99US-0145224.
 PR 26-JUL-1999; 99US-0145276.
 PR 27-JUL-1999; 99US-0145913.
 PR 27-JUL-1999; 99US-0145918.
 PR 27-JUL-1999; 99US-0145919.
 PR 28-JUL-1999; 99US-0145951.
 PR 02-AUG-1999; 99US-0146386.
 PR 02-AUG-1999; 99US-0146388.
 PR 03-AUG-1999; 99US-0146389.
 PR 04-AUG-1999; 99US-0147204.
 PR 05-AUG-1999; 99US-0147202.
 PR 05-AUG-1999; 99US-0147200.
 PR 06-AUG-1999; 99US-0147203.
 PR 06-AUG-1999; 99US-0147416.
 PR 09-AUG-1999; 99US-0147493.
 PR 09-AUG-1999; 99US-0147935.
 PR 10-AUG-1999; 99US-0148171.
 PR 11-AUG-1999; 99US-0148319.
 PR 12-AUG-1999; 99US-0148341.
 PR 13-AUG-1999; 99US-0148564.
 PR 13-AUG-1999; 99US-0148684.
 PR 16-AUG-1999; 99US-0149368.
 PR 17-AUG-1999; 99US-0149375.
 PR 18-AUG-1999; 99US-0149426.
 PR 18-AUG-1999; 99US-0149722.
 PR 20-AUG-1999; 99US-0149723.
 PR 23-AUG-1999; 99US-0149902.
 PR 23-AUG-1999; 99US-0149930.
 PR 25-AUG-1999; 99US-0150556.
 PR 26-AUG-1999; 99US-0150884.
 PR 27-AUG-1999; 99US-0151065.
 PR 27-AUG-1999; 99US-0151066.
 PR 27-AUG-1999; 99US-0151080.
 PR 31-AUG-1999; 99US-0151303.
 PR 31-AUG-1999; 99US-0151484.

Query Match 32.7%; Score 512.5; DB 21; Length 276;

Best Local Similarity 42.5%; Pred. No. 1; ge-45; Matches 113; Conservative 52; Mismatches 90; Indels 11; Gaps 4;

Qy 3 VGFIGAGQLAYRF---TAAGJLSAHKTAESSEMNLPTVSRLKMQVNLTNSKETVHK 58
 Db 13 vgfifagqkmaesiargvavsgvppnr-ictavhsnlrrdvyfesfgvnvfstseevke 71
 Qy 59 SDVFLFLAVKPHILPFILDEBIGDVQARHIVVSCAAGYVISSEKKLMAFQPAKPVIRCM 118
 Db 72 sdvififsvqpvkavtelskkskntlvsgaqiklndqe---wsqgdrfivmp 127
 Qy 119 NTPVWQEGATVYVMTGTHALVEDQOLQOLMSSVGFCTEVEDPLIDAVTGSSGSPAYF 178
 Db 128 ntpavgeeaavmvlsgtgeedgavamifgavgkilkadkmfdvtgjsgspayf 187
 Qy 179 MALO--ADCGVKGMLPRLRAIQLGAQALGAKMLDQSEOHPCOLKNCVSGGATHAL 236
 Db 188 laiceladgvgvaqprelalslasqtvlgatmvsktgkhpvgkvadvtspqgttay 247
 Qy 237 HFLESGGFSLLINAVEASCIRREL 262
 Db 248 helekgfsfratlmnavvaakrsrel 273

AAG05979
 ID AAG06979 standard; Protein; 256 AA.
 XX
 AC
 XX
 DT 17-OCT-2000 (first entry)
 XX
 DE Arabidopsis thaliana protein fragment SEQ ID NO: 3955.
 XX
 KW Protein identification; signal transduction pathway; metabolic pathway;
 KW hybridisation assay; genetic mapping; gene expression control; promoter;
 KW termination sequence.
 XX
 OS Arabidopsis thaliana.
 XX
 PN EP1033405-A2.
 XX
 PD 06-SEP-2000.
 XX
 PF 25-FEB-2000; 2000EP-0301439.
 XX
 PR 25-FEB-1999; 990US-0121825.
 PR 05-MAR-1999; 990US-0123180.
 PR 09-MAR-1999; 990US-0123548.
 PR 23-MAR-1999; 990US-0125788.
 PR 25-MAR-1999; 990US-0126264.
 PR 29-MAR-1999; 990US-0126785.
 PR 01-APR-1999; 990US-0127462.
 PR 06-APR-1999; 990US-0128234.
 PR 08-APR-1999; 990US-0128714.
 PR 16-APR-1999; 990US-0129845.
 PR 19-APR-1999; 990US-0130077.
 PR 21-APR-1999; 990US-0130449.
 PR 23-APR-1999; 990US-0130510.
 PR 23-APR-1999; 990US-0130891.
 PR 28-APR-1999; 990US-0131449.
 PR 30-APR-1999; 990US-0132048.
 PR 04-MAY-1999; 990US-0132484.
 PR 05-MAY-1999; 990US-0132485.
 PR 06-MAY-1999; 990US-0132486.
 PR 06-MAY-1999; 990US-0132487.
 PR 07-MAY-1999; 990US-0132863.
 PR 11-MAY-1999; 990US-0134255.
 PR 14-MAY-1999; 990US-0134219.
 PR 14-MAY-1999; 990US-0134221.
 PR 14-MAY-1999; 990US-0134370.
 PR 18-MAY-1999; 990US-0134768.
 PR 20-MAY-1999; 990US-0135414.
 PR 21-MAY-1999; 990US-0135353.
 PR 24-MAY-1999; 990US-0135629.
 PR 25-MAY-1999; 990US-0136021.
 PR 27-MAY-1999; 990US-0136392.
 PR 28-MAY-1999; 990US-0136782.
 PR 01-JUN-1999; 990US-0137222.
 PR 03-JUN-1999; 990US-0137528.
 PR 07-JUN-1999; 990US-0137724.
 PR 08-JUN-1999; 990US-0138094.
 PR 10-JUN-1999; 990US-0138540.
 PR 14-JUN-1999; 990US-0139119.
 PR 16-JUN-1999; 990US-0139452.
 PR 16-JUN-1999; 990US-0139453.
 PR 17-JUN-1999; 990US-0139492.
 PR 18-JUN-1999; 990US-0139497.
 PR 18-JUN-1999; 990US-0139455.
 PR 18-JUN-1999; 990US-0139456.
 PR 18-JUN-1999; 990US-0139457.
 PR 18-JUN-1999; 990US-0139458.
 PR 18-JUN-1999; 990US-0139459.
 PR 18-JUN-1999; 990US-0139817.
 PR 18-JUN-1999; 990US-0139839.
 PR 18-JUN-1999; 990US-0140353.
 PR 18-JUN-1999; 990US-0140354.
 PR 18-JUN-1999; 990US-0140595.
 PR 18-JUN-1999; 990US-0140763.
 PR 21-JUN-1999; 990US-0140891.
 PR 22-JUN-1999; 990US-0141287.
 PR 23-JUN-1999; 990US-0141842.
 PR 24-JUN-1999; 990US-0142154.
 PR 01-JUL-1999; 990US-0142205.
 PR 02-JUL-1999; 990US-0142390.
 PR 06-JUL-1999; 990US-0142390.
 PR 08-JUL-1999; 990US-0142390.
 PR 09-JUL-1999; 990US-0142390.
 PR 12-JUL-1999; 990US-0142377.
 PR 13-JUL-1999; 990US-0143542.
 PR 14-JUL-1999; 990US-0144624.
 PR 15-JUL-1999; 990US-0144632.
 PR 16-JUL-1999; 990US-0144085.
 PR 16-JUL-1999; 990US-0144086.
 PR 19-JUL-1999; 990US-0144325.
 PR 19-JUL-1999; 990US-0144331.
 PR 19-JUL-1999; 990US-0144332.
 PR 19-JUL-1999; 990US-0144333.
 PR 19-JUL-1999; 990US-0144334.
 PR 19-JUL-1999; 990US-0144335.
 PR 20-JUL-1999; 990US-0144352.
 PR 20-JUL-1999; 990US-0144632.
 PR 20-JUL-1999; 990US-0144684.
 PR 21-JUL-1999; 990US-0144814.
 PR 21-JUL-1999; 990US-0144814.
 PR 21-JUL-1999; 990US-0145088.
 PR 22-JUL-1999; 990US-0145088.
 PR 22-JUL-1999; 990US-0145087.
 PR 22-JUL-1999; 990US-0145089.
 PR 20-JUL-1999; 990US-0145199.
 PR 21-JUL-1999; 990US-0145192.
 PR 21-JUL-1999; 990US-0145192.
 PR 23-JUL-1999; 990US-0145218.
 PR 23-JUL-1999; 990US-0145224.
 PR 26-JUL-1999; 990US-0145276.
 PR 27-JUL-1999; 990US-0145913.
 PR 27-JUL-1999; 990US-0145918.
 PR 27-JUL-1999; 990US-0145918.
 PR 27-JUL-1999; 990US-0145919.
 PR 28-JUL-1999; 990US-0145951.
 PR 02-AUG-1999; 990US-0145386.
 PR 02-AUG-1999; 990US-0145388.
 PR 02-AUG-1999; 990US-0145389.
 PR 03-AUG-1999; 990US-0145038.
 PR 04-AUG-1999; 990US-0145204.
 PR 04-AUG-1999; 990US-0145302.
 PR 05-AUG-1999; 990US-0145192.
 PR 05-AUG-1999; 990US-0145260.
 PR 06-AUG-1999; 990US-0145303.
 PR 06-AUG-1999; 990US-0145416.
 PR 09-AUG-1999; 990US-0145493.
 PR 09-AUG-1999; 990US-0145935.
 PR 10-AUG-1999; 990US-0145171.
 PR 11-AUG-1999; 990US-0145319.
 PR 12-AUG-1999; 990US-0145341.
 PR 13-AUG-1999; 990US-0145656.
 PR 13-AUG-1999; 990US-0145684.
 PR 16-AUG-1999; 990US-0145368.
 PR 17-AUG-1999; 990US-0145175.
 PR 18-AUG-1999; 990US-0145426.
 PR 20-AUG-1999; 990US-0145722.
 PR 20-AUG-1999; 990US-0145923.
 PR 20-AUG-1999; 990US-0149929.

PR	23-AUG-1999;	99US-0149902.	Db	: : : : : : :
PR	25-AUG-1999;	99US-0149930.	Db	177 gvaagqprealstlasqtvigaatmvsktgkhpgvkladvtspqgtiagvheleksfr 236
PR	26-AUG-1999;	99US-0150566.	Qy	246 SLLINWVEASCIRTEL 262
PR	27-AUG-1999;	99US-015065.	Db	: : :
PR	27-AUG-1999;	99US-015066.	Db	237 atlmavaavaakrsrel 253
PR	27-AUG-1999;	99US-015080.		
PR	31-AUG-1999;	99US-0151303.		
PR	01-SEP-1999;	99US-0151438.		
PR	07-SEP-1999;	99US-0152363.		
PR	10-SEP-1999;	99US-0152070.		
PR	13-SEP-1999;	99US-0153758.		
PR	15-SEP-1999;	99US-0154018.		
PR	16-SEP-1999;	99US-0154039.		
PR	20-SEP-1999;	99US-0154779.		
PR	22-SEP-1999;	99US-0155139.		
PR	23-SEP-1999;	99US-0155486.		
PR	24-SEP-1999;	99US-0155659.		
PR	28-SEP-1999;	99US-0156458.		
PR	29-SEP-1999;	99US-0156596.		
PR	04-OCT-1999;	99US-0157117.		
PR	05-OCT-1999;	99US-0157753.		
PR	06-OCT-1999;	99US-0157865.		
PR	07-OCT-1999;	99US-0158029.		
PR	08-OCT-1999;	99US-0158232.		
PR	12-OCT-1999;	99US-0158369.		
PR	13-OCT-1999;	99US-0159293.		
PR	13-OCT-1999;	99US-0159294.		
PR	14-OCT-1999;	99US-0159295.		
PR	14-OCT-1999;	99US-0159299.		
PR	14-OCT-1999;	99US-0159331.		
PR	14-OCT-1999;	99US-0159337.		
PR	14-OCT-1999;	99US-0159338.		
PR	18-OCT-1999;	99US-0159384.		
PR	21-OCT-1999;	99US-016041.		
PR	21-OCT-1999;	99US-0160567.		
PR	21-OCT-1999;	99US-0160768.		
PR	21-OCT-1999;	99US-0160770.		
PR	21-OCT-1999;	99US-0160814.		
PR	22-OCT-1999;	99US-0160880.		
PR	22-OCT-1999;	99US-0160981.		
PR	22-OCT-1999;	99US-0160989.		
PR	25-OCT-1999;	99US-0161404.		
PR	25-OCT-1999;	99US-0161405.		
PR	25-OCT-1999;	99US-0161406.		
PR	26-OCT-1999;	99US-0161539.		
PR	26-OCT-1999;	99US-0161360.		
PR	26-OCT-1999;	99US-0161361.		
PR	28-OCT-1999;	99US-0161920.		
PR	28-OCT-1999;	99US-0161992.		
PR	28-OCT-1999;	99US-0161993.		
PR	29-OCT-1999;	99US-0162142.		
Qy	8 AGOLAYRETAGILSAHKIIASSPENNLPLT/SLRKMGNULTRSNKETVKKHSDVLF/FLAVK	67		
Qy	2 aesiargvavsgvlpnnr-ictavhnslnrrdvsfesgvnfvstseevkkesdvvifsk	60		
Db	1 :: : : : : : : : : : : : : :			
Db	68 PHIPFILDEGADGVARHIVSCAGVTTSSVERKLMAFQAPKVKIRMTNTPVVWBG	127		
Db	61 pqvvkavtekskisknknllvvagikindlqe---wsqqrifrmppnppaaagea	116		
Qy	128 AYVYAGCTHALVEDGOLLEQIMSSVGCTEVEDLDIAYVGLSGSGPAPAFMALD--ADG	185		
Qy	117 aevmsigtgateedgavamifgavqkilkadekindfadtgvlgsgspayifalealadg	176		
Qy	186 GVKMGFLPRLAOLGQAQALGAKMILDSBQHPCQKLDKNCSPGGATHALPHLESSGFR	245		
Qy	Query Match 31.1%; Score 487.5; DB 21; Length 256;			
Matches	best Local Similarity 41.6%; Pred. No. 8.1e-44; Mismatches 92; Indels 7; Gaps 3;			
Qy	8 AGOLAYRETAGILSAHKIIASSPENNLPLT/SLRKMGNULTRSNKETVKKHSDVLF/FLAVK	67		
Qy	1 :: : : : : : : : : : : : : :			
Db	2 aesiargvavsgvlpnnr-ictavhnslnrrdvsfesgvnfvstseevkkesdvvifsk	60		
Db	68 PHIPFILDEGADGVARHIVSCAGVTTSSVERKLMAFQAPKVKIRMTNTPVVWBG	127		
Db	61 pqvvkavtekskisknknllvvagikindlqe---wsqqrifrmppnppaaagea	116		
Qy	3 VGRIGACOLAYRFTAGILSAHKIIASSPENNLPLT--/SLRKMGNULTRSNKETVKKHSDVLF/FLAVK	58		
Db	5 iffigcgnmg kallgigllasqgvlpqggvwytpspdkvaalndqfginaaesaqeqa	63		
Qy	59 SDVLF/FLAVKPHIPFILDEGADGVARHIVSCAGVTTSSVERKLMAFQAPKVKIRMT	118		
Db	64 adiifavkkgimikivseitsisinkdsvlaagvtldqar--alghdkiklram	120		
Qy	Sequence 269 AA;			
Qy	Query Match 25.7%; Score 402.5; DB 22; Length 269;			
Matches	best Local Similarity 35.5%; Pred. No. 1.1e-34; Mismatches 98; Conservative 52; Mismatches 107; Indels 19; Gaps 6;			
Qy	3 VGRIGACOLAYRFTAGILSAHKIIASSPENNLPLT--/SLRKMGNULTRSNKETVKKHSDVLF/FLAVK	58		
Db	5 iffigcgnmg kallgigllasqgvlpqggvwytpspdkvaalndqfginaaesaqeqa	63		
Qy	59 SDVLF/FLAVKPHIPFILDEGADGVARHIVSCAGVTTSSVERKLMAFQAPKVKIRMT	118		
Db	64 adiifavkkgimikivseitsisinkdsvlaagvtldqar--alghdkiklram	120		

PR	06-AUG-1999;	99US-0147416.
PR	09-AUG-1999;	99US-0147493.
PR	10-AUG-1999;	99US-0148171.
PR	11-AUG-1999;	99US-0148319.
PR	12-AUG-1999;	99US-0148341.
PR	13-AUG-1999;	99US-0148565.
PR	13-AUG-1999;	99US-0148684.
PR	16-AUG-1999;	99US-0149368.
PR	17-AUG-1999;	99US-0149175.
PR	18-AUG-1999;	99US-0149426.
PR	20-AUG-1999;	99US-0149722.
PR	20-AUG-1999;	99US-0149723.
PR	20-AUG-1999;	99US-0149929.
PR	23-AUG-1999;	99US-0149902.
PR	23-AUG-1999;	99US-0149930.
PR	25-AUG-1999;	99US-0150566.
PR	26-AUG-1999;	99US-0150884.
PR	27-AUG-1999;	99US-0151055.
PR	27-AUG-1999;	99US-0151066.
PR	27-AUG-1999;	99US-0151080.
PR	30-AUG-1999;	99US-0151303.
PR	31-AUG-1999;	99US-0151438.
PR	01-SEP-1999;	99US-0151930.
PR	07-SEP-1999;	99US-0152363.
PR	10-SEP-1999;	99US-0153070.
PR	13-SEP-1999;	99US-0153158.
PR	15-SEP-1999;	99US-0154018.
PR	16-SEP-1999;	99US-0154039.
PR	20-SEP-1999;	99US-0154779.
PR	22-SEP-1999;	99US-0155139.
PR	23-SEP-1999;	99US-0155486.
PR	24-SEP-1999;	99US-0155659.
PR	28-SEP-1999;	99US-0156558.
PR	29-SEP-1999;	99US-0156996.
PR	04-OCT-1999;	99US-0157117.
PR	05-OCT-1999;	99US-0157753.
PR	06-OCT-1999;	99US-0157855.
PR	07-OCT-1999;	99US-0158029.
PR	08-OCT-1999;	99US-0158322.
PR	12-OCT-1999;	99US-0158659.
PR	13-OCT-1999;	99US-0159233.
PR	13-OCT-1999;	99US-0159394.
PR	13-OCT-1999;	99US-015995.
PR	14-OCT-1999;	99US-0159329.
PR	14-OCT-1999;	99US-0159330.
PR	14-OCT-1999;	99US-0159331.
PR	14-OCT-1999;	99US-0159637.
PR	14-OCT-1999;	99US-0159938.
PR	18-OCT-1999;	99US-0159884.
PR	21-OCT-1999;	99US-0160414.
PR	21-OCT-1999;	99US-0160667.
PR	21-OCT-1999;	99US-0160768.
PR	21-OCT-1999;	99US-0160770.
PR	21-OCT-1999;	99US-0160814.
PR	21-OCT-1999;	99US-0160815.
PR	22-OCT-1999;	99US-0160880.
PR	22-OCT-1999;	99US-0160881.
PR	22-OCT-1999;	99US-0160889.
PR	25-OCT-1999;	99US-0161004.
PR	25-OCT-1999;	99US-0161005.
PR	25-OCT-1999;	99US-0161406.
PR	26-OCT-1999;	99US-0161559.
PR	26-OCT-1999;	99US-0161650.
PR	28-OCT-1999;	99US-0161920.
PR	28-OCT-1999;	99US-0161992.
PR	28-OCT-1999;	99US-0161993.
PR	29-OCT-1999;	99US-0162142.

Query Match 23.7%; Score 372; DB 21; Length 7
 Best Local Similarity 52.7%; Pred. No. 8.7e-32;
 Matches 78; Conservative 23; Mismatches 45; Indels

1

Query Match Similarity: 23.6%; Score: 370; DB: 22; Length: 23.8%; pred: 35a-31;

270:

RESULT	15
AAGB1123	
ID	AAGB1123 standard; Protein; 295 AA.
XX	
AC	AAGB1123;
XX	
DT	04-SEP-2001 (first entry)
XX	
DE	Mycobacterium tuberculosis potential drug target protein SEQ ID 174.
XX	
Drug target; growth; organism viability; characterisation	
KW	

This invention relates to a method for identifying a nucleotide or polypeptide sequence that may be a drug target, or essential for growth or viability of an organism. Polynucleotide sequences AAG81947 - AAG81960 represent DNA encoding proteins AAG8196 - AAG8121, *Mycobacterium tuberculosis* proteins which are potential drug targets. The DNA and protein sequences are used to illustrate the method of the invention. The method involves providing an unknown nucleotide or polypeptide sequence and comparing it to a number of sequences along with at least one algorithm capable of analysing a functional relationship between nucleotide and polypeptide sequences. The method is useful for characterising the function of nucleic acids and polypeptides that may be useful as a target for a drug or essential for the growth or viability of an organism.

Query Match		Score	DB	Length
Best	Local Similarity	19.58;	Score	306.5;
Matches	34.38;	DB	22;	Length 295;
Oy	48	LTRSNKENTVKHSDYLELAVKP-HIIPILDEGADY---QARHITWVSCAGVTISSE	91;	Mismatches 86;
Db	54	lvtasaadavenatfvvvavkpkadvepiyviadiananaaendaesaqyfvttvvagitiayfe	101	Indels 35; Gaps

